Business Dynamic Statistics of Innovative Firms*

Nathan Goldschlag† Elisabeth Perlman‡

December 21, 2017

Abstract

A key driver of economic growth is the reallocation of resources from low to high productivity activities. Innovation plays an important role in this regard by introducing new products, services, and business methods that ultimately lead to increased productivity and rising living standards. Traditional measures of innovation, particularly those based on aggregate inputs, are increasingly unable to capture the breadth and depth of innovation in modern economies. In this paper, we describe an effort at the US Census Bureau, the Business Dynamics Statistics of Innovative Firms (BDS-IF) project, which aims to address these challenges by extending the Business Dynamics Statistics data to include new measures of innovative activity. The BDS-IF project will produce measures of firm, establishment, and employment flows by firm age, firm size, and industry for the subset of firms engaged in activities related to innovation. These activities include patenting and trademarking, the employment of STEM workers, and R&D expenditures. The flexibility of the underlying data infrastructure allows this measurement agenda to be extended to include copyright activity, management practices, and high growth firms.

Keywords: firm dynamics, innovation, Longitudinal Business Database, Business Dynamics Statistics

*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 Shawn Klimek and Joe Staudt for helpful comments and suggestions.

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

‡U.S. Census Bureau, `elisabeth.perlman@census.gov`

1 Introduction

One of the most fundamental characteristics of modern market economies is continual process of change and renewal by which old products and methods give way to the new (Schumpeter, 1942). Innovation is an important mechanism of change in the economy, facilitating the introduction of new products, services, and business methods (Romer, 1990; Aghion and Howitt, 1992). Measuring innovative activity in the economy has proven difficult, as the key inputs and outputs of innovative activity are often novel, complex, and fundamentally hard to observe and measure (Smith, 2005). Traditional measures of innovation, particularly those based on aggregate inputs and outputs of innovative processes, are increasingly unable to capture the breadth and depth of economic change in modern economies (OECD, 2007, 2010). Therefore, there is significant interest among both researchers and policy makers in the development of new statistics that describe innovative activity in the economy and its evolution over time. Innovative processes are intimately linked with firm dynamics; the entry, exit, growth, and contraction of businesses reallocates labor and capital away from low productivity firms to high productivity ones (Decker, Haltiwanger, Jarmin, and Miranda, 2016, 2017; Acemoglu, Akcigit, Bloom, and Kerr, 2013). This means that the Business Dynamics Statistics (BDS) program at the U.S. Census Bureau is uniquely suited to investigate both the characteristics of innovative firms and their dynamics over time.

In this paper, we describe the Business Dynamics Statistics of Innovative Firms (BDS-IF), a project at the Census Bureau, which aims to produce new public-use statistics extending the BDS data to focus on firms engaged in innovative activities. The BDS program, built upon the Longitudinal Business Database (LBD), produces statistics that characterizes firm, establishment, and employment dynamics of US businesses across a number of dimensions including industry, geography, firm size, and firm age (Jarmin and Miranda, 2002).[1] The

[1]Establishments, in these data, represent a single physical location where business is conducted or services or industrial operations are performed. Firms are defined as business organizations consisting of one or more domestic establishments grouped under common ownership or control. In the case of single-unit businesses, the firm and establishment are identical. See https://www.census.gov/ces/dataproducts/bds/ for details. A major effort is currently underway to improve the underlying longitudinal linkages found in the LBD. See

BDS-IF project builds upon these data by measuring the subset of firms engaged in activities related to innovation. In this paper we describe the statistics that track the dynamics of firms engaged in patenting and trademarking and in industries that employ a disproportionate share of STEM workers. The flexibility of the underlying data infrastructure allows this measurement agenda to be extended further. These extensions might include firms that own copyrights, firms that report R&D expenditures, and firms with particular types of management practices.

The economic measurement community and statistical agencies in particular have expended considerable effort to capture specific inputs of innovative activity in official statistics. For several reasons, these efforts have often failed to produce a comprehensive view of innovative activity in the economy (OECD, 2010). In many cases these efforts develop measures of innovation that are not flexible enough to capture what often appears a moving target in the data. Then BDS-IF project, by focusing on a wide variety of measures, is better able to capture the various dimensions of innovation, which often vary significantly across industries.

One of the most common measures of innovative activity is R&D expenditures collected from surveys of businesses. For example, the Survey of Industrial Research and Development (SIRD) and its successor the Business Research and Development and Innovation Survey (BRDIS) , sponsored by the National Science Foundation (NSF) and collected by the Census Bureau, ask a sample of U.S. businesses about their R&D activities (Foster, Grim, and Zolas, 2016).[2] While the BRDIS frame is known to be dominated by the largest R&D performing firms, the planned BRDIS-M (micro) survey will improve the measurement of smaller R&D performing firms by over-sampling small firms. In addition to the R&D activities of private firms, we may also be interested in measuring the impacts of university-led research, which is often more basic in nature when compared to private R&D expenditures. Leveraging various

Stinson, White, and Lawrence (2017) for details.

[2]R&D statistics from BRDIS feature prominently in the NSF's Science and Engineering indicators publication, a biennial report to Congress provides a broad base of quantitative information about US science, engineering, and technology.

administrative data sources, researchers at the Census Bureau aim to directly measure the impacts of university-based R&D expenditures (Zolas et al., 2015). This work focuses on both the inputs and outputs of university-based research. The inputs in this framework include the observed labor and capital used on grant funded research projects and the outputs include research trained workers and PhD students, patents, and publications.

In contrast to R&D expenditures, which represent inputs to the innovation process, patents capture the outputs or results of innovative activity. A substantial literature has explored the use of patent statistics as indicators of innovative activity, paying close attention to types of things patent statistics capture and the types of things they do not (Griliches, 1998). More recently, the Census Bureau began a new survey, the Annual Survey of Entrepreneurs (ASE), which provides information on entrepreneurship, an important driver of reallocation and growth (Foster and Norman, 2016). The ASE focuses on young firms and the experiences of their owners. The 2014 reference year ASE included an innovation module, capturing information about a variety of innovative activities firms may be engaged in. Finally, another way we might identify segments of the economy that have undergone periods if intense innovative activity is by examining the patterns of firm, employment, and productivity dynamics (Foster, Grim, Haltiwanger, and Wolf, 2017). Much like inferring the presence of black holes by observing their impact on nearby matter in space, this approach uses the surge of entry, reallocation, and subsequent productivity growth to identify innovative activity in the economy.

Our approach in the BDS-IF project, reflective of the fact that innovation has many dimensions, will be to develop multiple measures of innovative activity and integrate them within a common framework. We use the BDS data infrastructure as a platform to measure the dynamics of innovative firms, which are identified based on both inputs (STEM employment, R&D spending) and outputs (patents, trademarks, copyrights) of the innovation process. As a starting point, these statistics will capture firm, establishment, and employment flows by firm age, firm size, and industry composition of innovative firms.[3] These measures can be further refined to provide additional detail specific to the type of innovative activity under consideration. For example, for patenting firms we might be interested in the entry and exit of firms that have been granted highly cited patents or patents in specific technology classes. For R&D performing firms, on the other hand, we might consider differences in employment flows of firms engaged in basic versus applied R&D activities. In the following sections, we will provide an overview of the active components of the BDS-IF project and the progress made in developing the measurement infrastructure necessary to measure the business dynamics of innovative firms.

The remainder of the paper is organized as follows: In Section 2 we describe a number of active and ongoing projects aimed at improving the measurement of innovation in the economy. We describe projects to characterize the dynamics of patenting firms, trademarking firms, and firms in the High Tech sector. In Section 3 we describe additional extensions and future work that could be done to capture additional dimensions of the innovative activity in the economy. Section 4 concludes.

2 Innovative Firms

There is no obvious, universally agreed upon definition of an innovative firm. There are many activities firms engage in that might signal the presence of innovative activity. As noted above, our initial focus will be on patenting, trademarking, and employing a high proportion of STEM workers.[4] Firms might also be innovative at one point in time, but not at another; it is not obvious when this switch occurs. For example, though we know the date a patent is granted, patent grants take years to acquire and the activities that led to the patented innovation may have occurred years earlier. In the following sections we

[3]Here we use the term "innovative firms" loosely, since which firms are designated as "innovative" will depend on the specific measure under consideration. For example, the Business Dynamics Statistics of Patenting Firms component will focus on firms that have been granted patents.

[4]Science, Technology, Engineering, and Mathematics. For additional information about BLS Standard Occupation Classification codes associated with STEM see https://www.bls.gov/soc/.

describe in more detail several efforts, projects at different stages of completion, to identify innovative firms, each of which feed into the BDS-IF measurement agenda and can be used to characterize a specific type of innovative activity in the economy.

2.1 Patenting Firms

In the United States, patents have long been granted to the "first and true inventor" of the innovation in question, thus making the patent record an obvious place to look for evidence of innovation.[5] Over 200 years of law have helped define the parameters for a patent grant, but the primary criteria of utility, novelty, and nonobviousness have been in place since the patent act of 1790.[6] Patent statistics, therefore, provide a useful window into innovative activity in the economy.

It is clear that not every innovation is patented, or even patentable, and not every patent represents an important innovation. However, despite well-known limitations, patents have proven to be a useful proxy when measuring innovative activity (Griliches, 1998; Hall, Jaffe, and Trajtenberg, 2001; Thompson and Fox-Kean, 2005).[7] One significant advantage of patent data is that it contains a wealth of information that characterizes the nature of each innovation. Patent grants create a limited-term monopoly, entitling the holder to bar others from using the patented invention. This means that a patent must specify what the invention is, as well as the boundaries defining what the invention is not. In service of this, patent documents include information on the inventor(s), such as name and location, the name and location of the firm to which the patent was assigned, detailed descriptions of the contents of the innovation, related innovations in the form of prior art, and classification

[5] The U.S. was one of the few counties in the world to adopt a first-to-invent system rather than a first-to-file system. The America Invents Act (2011) changed this; all patents filed on or after March 16, 2013 are covered under a first-to-file regime.

[6] An 1842 patent act created the category of Design, as supposed to Utility, patents. These patents protect attributes that are ornamental - though they must also be novel and nonobvious. Design patents account for about 8% of our database.

[7] Utility patents cover processes, machines, manufacture, or compositions of matter. The USPTO also issues design patents, which cover designs embodied in an article of manufacture, and plant patents, which protect asexually reproduced plants. For more details see https://www.uspto.gov/web/offices/ac/ido/oeip/taf/patdesc.htm.

information describing the type of innovation.

In order to characterize the types of firms engaged in patenting activity, several efforts have been made to link patent assignees to other firm-level data. Early examples include the link of patent assignees to businesses in Compustat in Cummins, Hall, Laderman, and Mundy (1985) and Hall, Jaffe, and Trajtenberg (2001). There were also several early efforts to link patent assignees to Census Bureau data in Kerr and Fu (2008) and Balasubramanian and Sivadasan (2010). These linkages provided a rich set of covariates for understanding the relationship between firm-level measures of R&D, innovation, and productivity. However, all of these linkage efforts relied on relatively imprecise fuzzy name and address string matching techniques, which suffer from the limited geographic detail available in the patent data. Graham, Grim, Islam, Marco, and Miranda (2015) (henceforth GGIMM) introduce a novel methodology that dramatically improves the number and quality of matches. GGIMM leverage information on both inventors and assignees to create robust longitudinal firm-level linkages between patent assignees and firms in the LBD.

The Business Dynamics of Patenting Firms (BDS-PF) component of the innovation project builds upon the matching methodologies introduced by GGIMM to create and maintain robust microdata firm-assignee linkages. The BDS-PF project then uses these linkages to measure the business dynamics of patenting firms. These types of statistics allow us to characterize the job creation and job destruction patterns associated with patenting firms across a number of characteristics such as firm age and firm size. The initial vintage of these linkages will be made available to qualified researchers on approved projects through the Federal Statistical Research Data Center (FSRDC) network. Below we provide a preview of the types of statistics that result from the firm-assignee linkages developed as part of the BDS-PF project.

Consistent with findings in GGIMM, and more broadly within the innovation literature, we find that patenting firms tend to be relatively larger and older.[8] Figure 1 shows the mean

[8]For example, see Scherer (1965), Cohen (2010), Acs and Audretsch (1991) and more recently Akcigit and Kerr (2017).

share of firms across firm age and firm size groups for 2000 to 2014.[9] In panel (b) of Figure 1 we see about a quarter of patenting firms each year are in the oldest left censored category, which in 2000 would include firms that were at least 24 years old.[10] Comparing panel (a) and (b), only about 3.8% of patenting firms are startups compared to 9.3% for all firms in the same time period. Looking across the age groups, the gap between the share of all firms versus the share of patenting firms is largest for the younger groups, falling steadily across older groups. Panel (c) and (d) of Figure 1 shows that, whereas the majority (76%) of all firms have fewer than 10 employees, only about 27% of patenting firms have fewer than 10 employees. The share of firms in the larger firm size groups is vanishingly small among all firms, but about a quarter of patenting firms have 250 or more employees. It is important to note that though few in number, these large firms account for a disproportionate share of total employment – firms with 250 or more employees account for 57% of total employment and 98% of employment among patenting firms on average between 2000 and 2014.

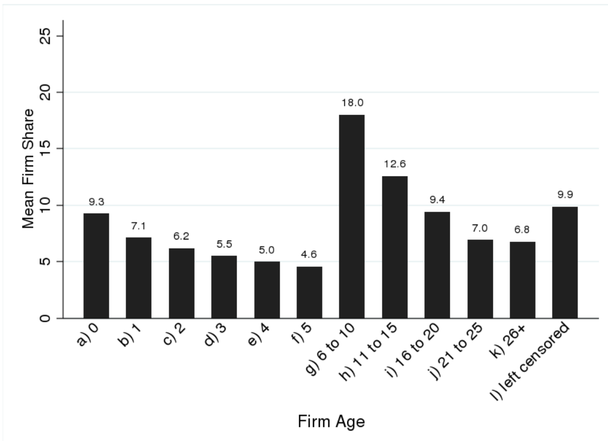
Patenting activity is over represented among larger and older groups of firms. However, comparing the distribution across firm age and size categories does not provide a sense of the relative propensity to patent within size and age groups. Being granted a patent is a very rare event for firms in our data–on average only about 0.3% of firms receive a patent grant. Though a tiny fraction of firms, patenting firms are large–they accounted for 23% of total employment.

The left panel of Figure 2 shows that patenting firms account for over 40% of all employment among the oldest, left censored firms. Patenting firms also account for a sizable share of employment among firms aged 21 to 25 years and 26+ years (14% and 19% respectively). In contrast, patenting firms account for only about 0.6% of employment among startups. By firm size, in the right panel of Figure 2, we see that 64% of employment among the largest firms is in firms with patent grants. The share of employment within each firm size group

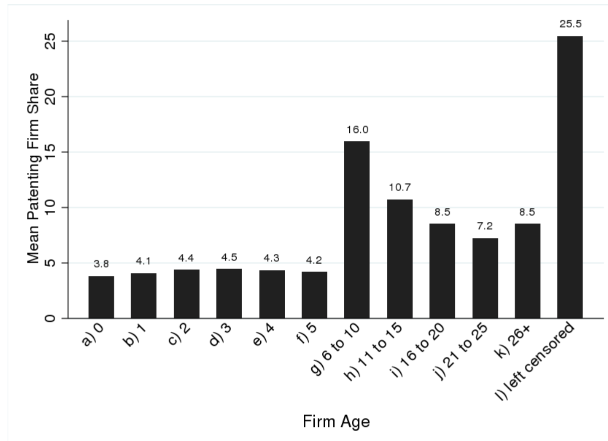
[9] That is, the share is computed within each year, and we present the mean share across years.

[10] Due to data limitations a firms' age is considered left censored if it was founded before 1976. Thus the oldest observable age increases with time.

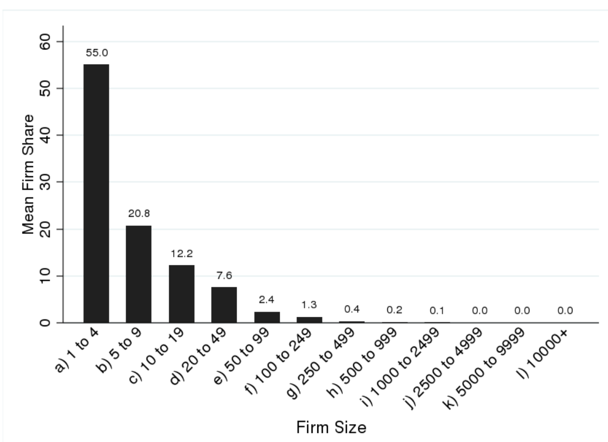
Figure 1: Patenting Firm Share by Firm Age (a,b) and Firm Size (c,d)



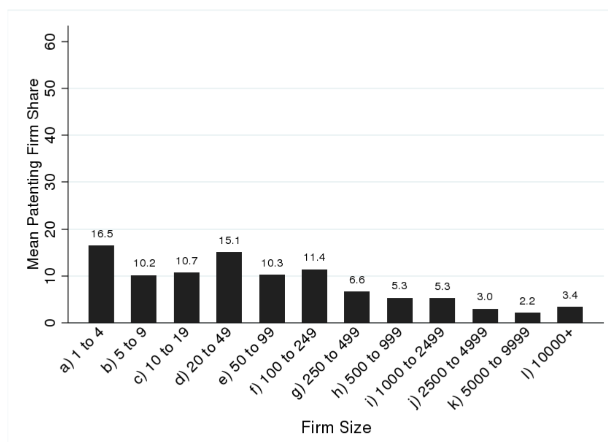
(a) All Firms, by Firm Age



(b) Patenting Firms, by Firm Age



(c) All Firms, by Firm Size

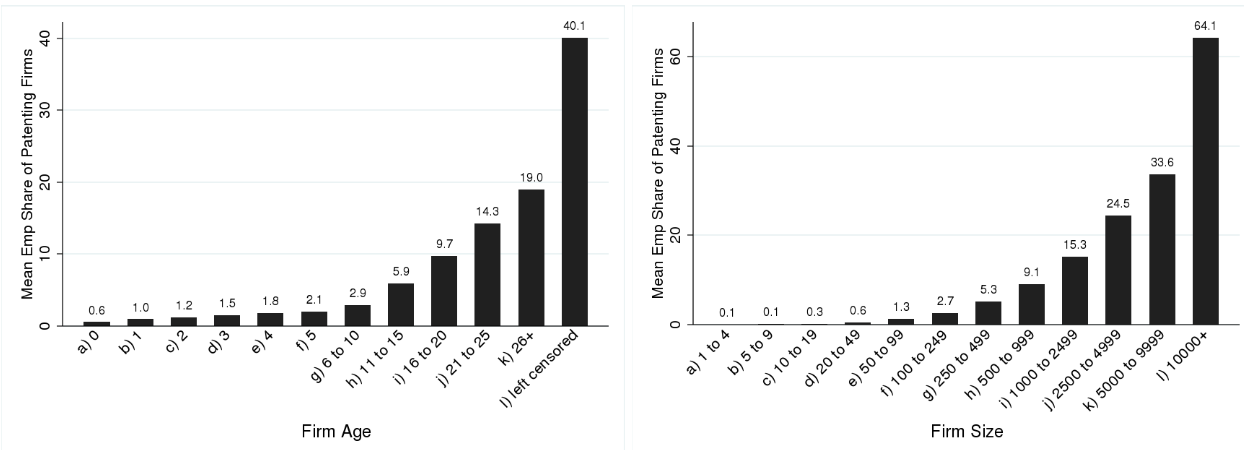


(d) Patenting Firms, by Firm Size

Source: Longitudinal Business Database, BDS-PF Crosswalk, author's calculations.
Notes: Mean shares within groups calculated for each year 2000-2014, then averaged across years. Contemporaneous patent grants used to identify patenting firms. Shares approximately sum to 100.

in patenting firms falls monotonically with firm size but remains sizable for all of the largest size groups. Moreover, it is the largest firms within each firm size group that are granted patents, causing the patenting employment share to be consistently higher than the patenting firm share across firm size categories. These figures suggest that even though patenting is relatively rare, the vast majority of activity among the largest firms occurs within firms that are granted patents.

Figure 2: Within Firm Age (left) and Firm Size (right) Patenting Employment Share



Source: Longitudinal Business Database, BDS-PF Crosswalk, author's calculations.
Notes: Mean shares within groups calculated for each year 2000-2014, then averaged across years. Contemporaneous patent grants used to identify patenting firms. Note different y-axis scale for left and right panels.

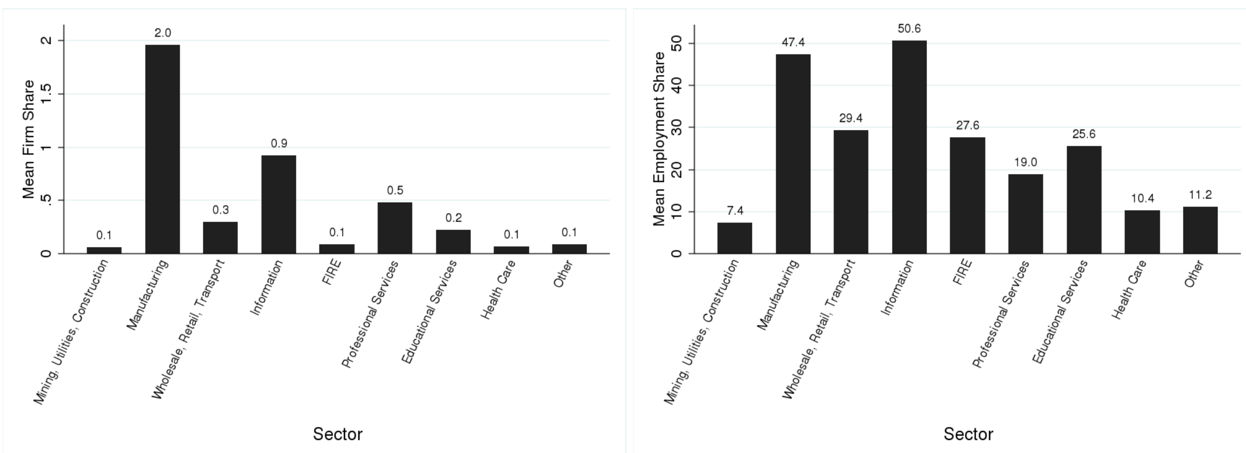
Consistent with findings in the innovation literature, we find substantial heterogeneity in the propensity to patent across sectors. Figure 3 shows the average firm share and employment share of patenting firms within industries.[11] The Manufacturing sector, often noted as having particular importance to innovation, has the highest percentage (2.0%) of firms granted patents, followed by Information (0.9%) and Professional Services (0.5%) sectors.[12]

[11]Sectors classified using NAICS codes. Mining, Utilities, Construction includes NAICS 21, 22, and 23, Manufacturing includes NAICS 31, 32, and 33, Wholesale, Retail, Transport includes NAICS 42, 44, 45, 48, and 49, Information includes NAICS 51, Finance, Insurance, and Real Estate (FIRE) includes NAICS 52, and 53, Professional Services includes NAICS 54, Educational Services includes NAICS 61, Health Care includes NAICS 62, and Other includes all other industries.

[12]See Locke and Wellhausen (2014) for details on the link between manufacturing production and innovation.

This ordering changes if we compare employment shares. By share of employment, Information leads with patenting firms accounting for 51% of employment in the sector, followed by Manufacturing (47%), and Wholesale, Retail, and Transport (29%). The difference between firm and employment shares is likely due to compositional differences across firm size categories across sectors, but regardless the pattern of patenting overwhelming occurring in very large firms remain.

Figure 3: Patenting Firm (left) and Employment (right) Share by Sector



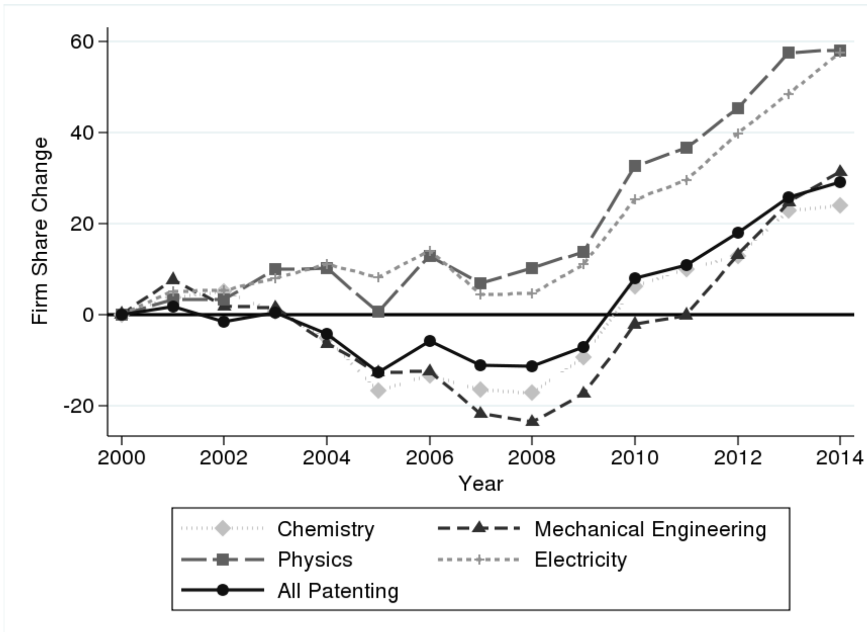
Source: Longitudinal Business Database, BDS-PF Crosswalk, author's calculations.
Notes: Mean shares within groups calculated for each year 2000-2014, then averaged across years. Sectors classified using Fort-Klimek NAICS industry codes Fort and Klimek (2016), see the text for the NAICS-Industry mapping used. Note different y-axis scale for left and right panels. Sector for multi-unit firms is assigned using employment weights at the establishment level.

In addition to illuminating firm characteristics of patenting firms, such as size, age, and sector, the underlying data infrastructure generated by the BDS-PF project also allows us to measure firm dynamics along dimensions native to the patent data. For example, as alluded to in previous sections, we can calculate measures of firm and employment flows by patent technology classes. Figure 4 shows the change in the share of firms receiving patent grants in different technology classes over time.[13] Note that the figure is indexed to firm share values in 2000 to show changes over time. We can see that the share of firms receiving patent grants

[13]Technology classes defined using CPC sections C "Chemistry; Metallurgy", F "Mechanical Engineering; Lighting ; Heating; Weapons; Blasting Engines or Pumps", G "Physics", and H "Electricity". For details see https://www.uspto.gov/web/patents/classification/cpc.html.

in Chemistry and Electricity has increased by about 50% since 2000. This compares to an overall increase in the share of firms with patent grants over the same period of about 16%.

Figure 4: Change in Firm Share Patenting in Technology Classes



Source: Longitudinal Business Database, BDS-PF Crosswalk, USPTO PatentsView database, author's calculations.
Notes: Firm share with patent grants in each year within technology classes indexed to 2000. Technology classes defined using CPC sections.

These results demonstrate the type of statistics that can be produced using the BDS-PF data infrastructure. In the future, the project aims to release a suite of statistics similar to those presented here, which we be useful and interesting to data users interested in innovation, patenting, and firm dynamics.

2.2 Trademarking Firms

Trademarks are a type of intellectual property that allows a business to protect a word, phrase, symbol, or design.[14] These marks signal, and were created to signal, valuable information about the origin of a product or service. In doing so, trademark protections facilitate

[14]The definition also includes color, smell and sound.

the creation of brand loyalty, reduce consumer search and switching costs, and lower the cost of introducing and marketing new products. Trademarks may provide a more direct measure of new product introductions, particularly in service and retail industries, than do R&D expenditures or patents. As such, trademarks are a valuable signal of innovative activity. In this section, we describe efforts to link information about trademarks to Census microdata assets.

The United States Patent and Trademark Office (USPTO) has made available trademark data that covers almost 7 million trademark registrations or applications filed for the period 1870-2015 (Graham, Hancock, Marco, and Myers, 2013).[15] These data include detailed information on trademark applications, registrations, commercial use, renewals, assignments, cancellations, and abandonments. One notable feature of the USPTO trademarks data is that they necessarily capture the transfer of trademarks among firms and could therefore be used to develop a comprehensive firm-level trademark portfolios (this is in contrast to patents, where reporting transfers is optional). Dinlersoz, Goldschlag, Myers, and Zolas (2017) (henceforth DGMZ) describe the creation of firm-level linkages of trademarking data to Census Bureau business data. These linkages allow the authors to document basic facts about what types of firms trademark as well as the timing of trademark activity within in the firm's life-cycle. DGMZ also use these linkages to provide the first look at the link between trademarking activity and firm outcomes including employment and revenue growth.

Here we summarize some key findings in DGMZ, paying particular attention to the use of these linkages to measure the business dynamics of innovative firms. One of the most basic contributions in DGMZ is the systematic documentation of the characteristics of trademarking firms. It is important to note that, because of their focus on first time trademarking activity, DGMZ restrict their analysis to firms born in 1976 or later. This restriction will skew the firm age distribution in early years towards young firms but will gradually become more representative over time. Table 1 shows the mean firm size and firm

[15] Federal trademarks were created in the US in 1870 but quickly declared unconstitutional. Subsequently, the 1881 trademark law was upheld.

age of firms that have at least one trademark ("With TM"), firms that file a TM for the first time ("First time TM"), and firms that do not have any TMs ("No TM"). From Table 1 we can see that first time trademarkers are relatively large and young compared to the set of firms that do not have any trademarks. Firms that have at least one trademark tend to be much larger and a bit older than firms with no trademarks.

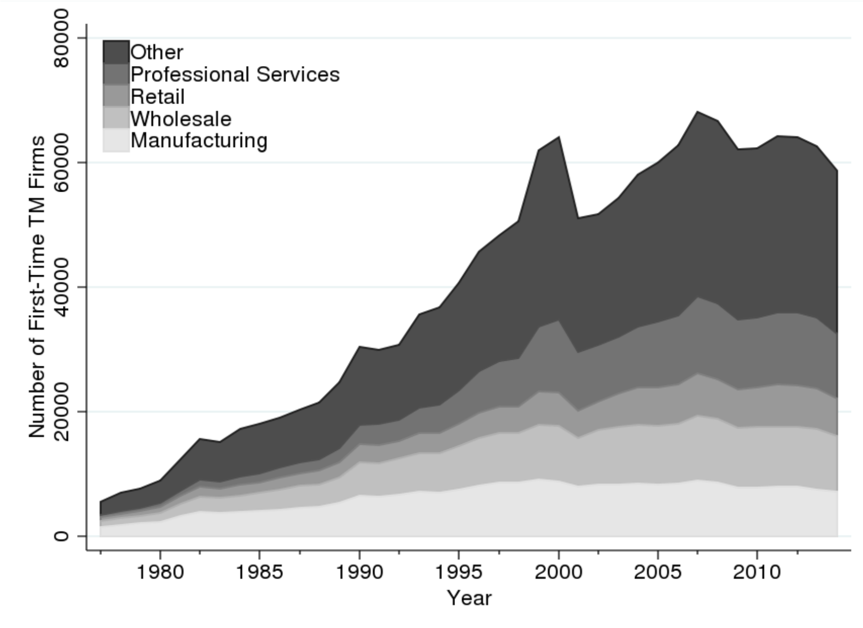
Table 1: Firm Size and Age of Trademarking Firms

	With TM		First Time TM		No TM	
Year	Mean Size	Mean Age	Mean Size	Mean Age	Mean Size	Mean Age
1997	100.3	7.8	76.1	4.0	9.5	6.4
2001	101.4	8.9	55.1	4.6	10.0	7.6
2005	89.7	10.3	48.0	4.9	9.5	8.2
2009	84.6	11.5	48.1	5.9	9.5	9.6
2013	84.1	13.9	57.7	8.3	9.7	10.8

Source: Dinlersoz, Goldschlag, Myers, and Zolas (2017)
Notes: Mean size is employment based and firm age is calculated based on payroll. Sample excludes "left censored" firms, or those born prior to 1977.

Figure 5 shows the distribution of first-time trademarking firms by sector over time. First-time trademarking activity grows for all sectors early in the time series in part due to the focus on firms born since 1976, which skews the sample towards young firms that are more likely to trademark for the first time. First-time trademarking activity in the "Other" and professional services sectors has grown substantially since the 1990s. First-time trademarking in the manufacturing sector, on the other hand, has declined over the same period. This decline is likely driven the be decline in manufacturing activity in the U.S. We also see a clear rise and fall of first-time trademarking activity around the dot-com boom and bust in the late 1990s driven primarily by trademarking activity among young firms (those less than 6 years old).

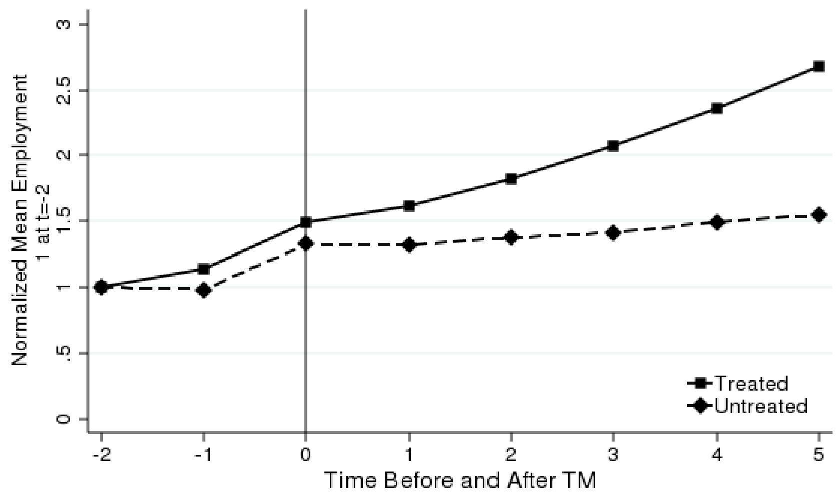
Figure 5: First-Time Trademarking by Sector



Source: Dinlersoz, Goldschlag, Myers, and Zolas (2017)
Notes: Count of first-time trademarking firms by sector by year.

Aside from a better understanding of the characteristics of trademarking firms we may also be interested in how trademarking impacts subsequent employment growth. Measuring this relationship is difficult because of the selection effects driving what firms seek trademark protection. Firms that anticipate growth may be more likely to pursue the protections offered by trademarks, contaminating any simple regression results examining the pre and post growth outcomes of trademarking firms. In order to assess the impacts of first-time trademarking activity and control for these selection effects DGMZ use propensity score matching to generate a control group. The propensity score model is estimated using observable characteristics including firm size, firm age, average payroll, multi-unit status, industry fixed effects, and prior year employment. The estimated probabilities are then used to select a set of non-treated (non-trademarking) firms that are substantively similar on observables to the treated (trademarking) group. Figure 6 shows the mean employment for firms that trademark for the first time in $t = 0$ versus the matched control group that do not. For comparison, both series are normalized to 1 at $t = -2$. In the year of first-time trademarking the trademarking firms are slightly larger than the control group that did not trademark and both show an upward trend from $t = -1$ to $t = 0$. The treated (trademarking) firms experience robust employment growth through $t+5$ while the control group grows much less. After five years the mean employment for firms that trademark is about 170% larger than two years prior to the filing. The control group, on the other hand, is about 50% larger after five years. All of the within year differences between the treated and control are statistically significant.

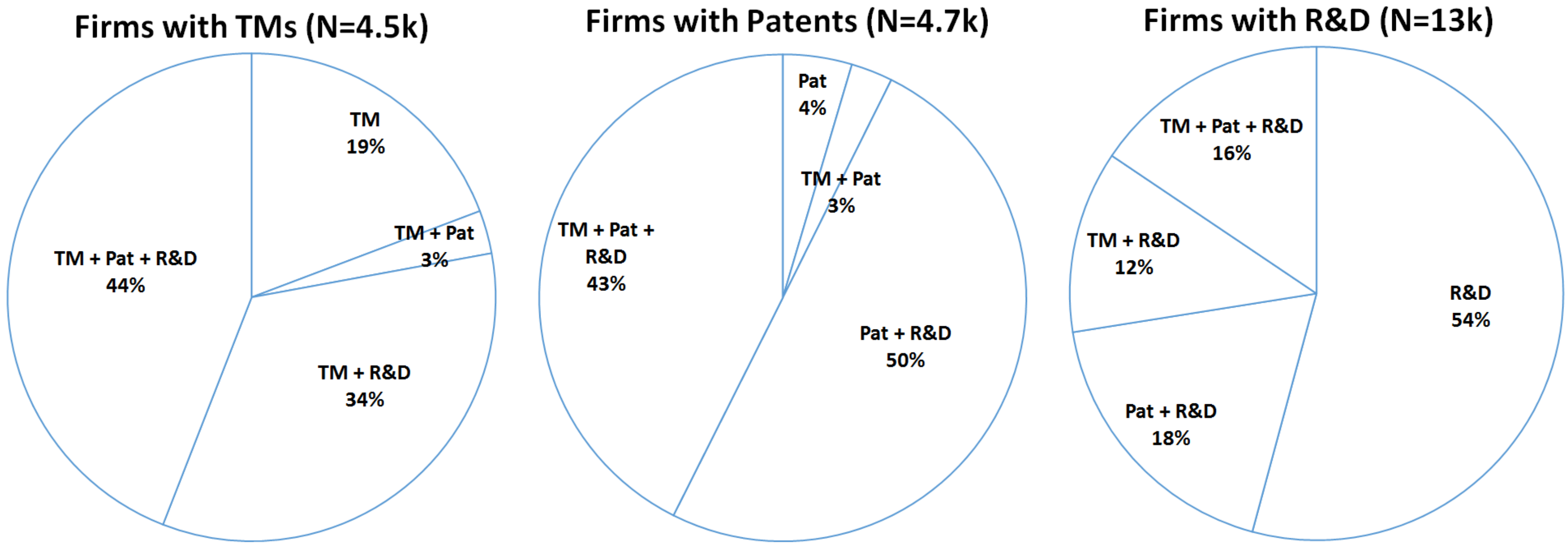
Figure 6: Impact of First-Time Trademarking on Employment



Source: Dinlersoz, Goldschlag, Myers, and Zolas (2017)
Notes: Average employment before and after first-time trademark filing, normalized to $t-2$, treated versus control. Control group identified using propensity score weights.

Finally, DGMZ unpack the relationship between trademarking and other markers of innovative activity such as patents and R&D expenditures. Figure 7 shows the overlap in trademarking, patenting, and R&D expenditures among firms in the BRDIS sample form in 2011. The far left panel shows to coincidence of trademarking, patenting, and R&D among firms with at least one trademark filing in their lifetime as of 2011 and are in the BRDIS sample. The middle panel shows the same for firms with at least one patent grant and the right panel shows firms with positive R&D expenditures at some point up to 2011. In the left panel, we can see that a large fraction (44%) of BRDIS firms with at least one trademark also have patents and report positive R&D expenditures. Moreover, 81% of firms included in BRDIS that have filed for at least one trademark also engage in patenting and/or R&D. Despite this, there is still a significant share (19%) only file for trademarks without being found with patents or R&D expenditures. In contrast, the middle panel shows that the overwhelming majority (93%) of BRIDS firms with at least one patent also report R&D expenditures. Almost half of these firms (46%) also filed for at least one trademark. Finally, in the right panel we see that a significant share (54%) of R&D performing firms do not have any trademarks or patents. About 16% of R&D performing firms, on the other hand, are also found to have patent grants and trademark filings. Overall, Figure 7 shows that trademarking not only correlates highly with other innovative activities such as patenting and R&D expenditures, but it also identifies a significant number of firms that only trademark. To the extent that trademarking is indicative of innovative activities different from those captured by patent and R&D measures, the wider net cast by the trademarking links help identify a larger set of innovative firms.

Figure 7: Trademarking, Patenting, and R&D Expenditures



Source: Dinlersoz, Goldschlag, Myers, and Zolas (2017)

Notes: Percentages based on unweighted counts of firms sampled in BRDIS that are found to have at least one trademark filing (left), at least one patent grant (middle), or positive R&D expenditures in 2011 or prior years.

2.3 High Tech Industries

Whereas patents and trademarks are outputs of innovative activity, another way to measure innovative activity in the economy is to focus in on a subset of industries we suspect to be creating and diffusing productivity enhancing technologies. The High Tech sector is one such set of industries, responsible for the creation of many new products and services that drive reallocation and growth.[16] Developing a better understanding of these types of firms has therefore become an important component the growing innovation and reallocation literature (Stoneman and Battisti, 2010; Acemoglu, Akcigit, Bloom, and Kerr, 2013; Decker, Haltiwanger, Jarmin, and Miranda, 2016). Goldschlag and Miranda (2016) (henceforth GM) describe the creation of new public use data products that measure the business dynamics of firms in High Tech industries.

Identifying which industries are High Tech is a non-trivial problem. GM review a number of different methodologies that have been proposed over the years to define the High Tech sector. Generally speaking, these methodologies focus on either the inputs or outputs of the innovation process. For example, one might identify High Tech industries based on patterns in R&D investments across industries. Alternatively, the concentration of employment in technology oriented occupations could be used to define High Tech. Ultimately, GM extend the Hecker (2005) methodology, identifying High Tech industries as those with relatively high concentrations of STEM employment. This method has several key advantages with respect to its application in creating Business Dynamics Statistics. First, it provides comprehensive coverage of industries outside of manufacturing. Second, the occupation data used to calculate STEM concentrations are readily available, which allows data users to replicate the exercise and allows the classification to be updated if necessary. Finally, industries identified as High Tech is robust over time and with respect to alternative STEM occupation classifi-

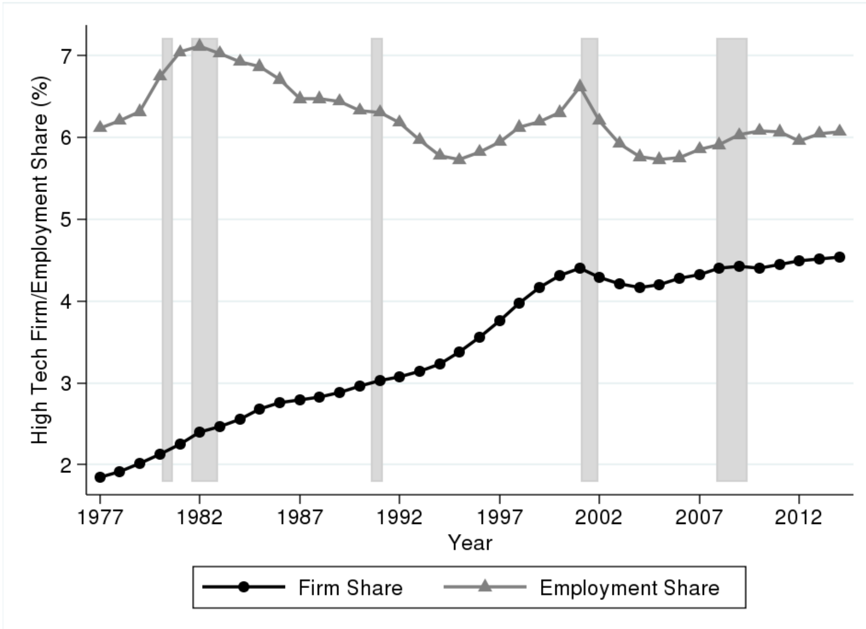
[16]The NSF published figures from the 2011 BRDIS that measure the share of businesses that introduce new goods or services between 2009-2011. These include estimates for 9 of the 15 4-digit NAICS industries classified as High Tech in Goldschlag and Miranda (2016). On average, 40% businesses in the covered High Tech industries introduced a new product or service compared to 9% for all industries and 22% for manufacturing. See https://www.nsf.gov/statistics/2016/nsf16308/overview.htm for details.

cations. The set of industries that employ the highest share of STEM workers do not tend to change over time.

The GM High Tech classification identifies 15 4-digit 2007 NAICS industries. The classification includes industries in mining, manufacturing, information, and professional services. There are several subtle points of interpretation, highlighted in GM, that must be kept in mind. First, the classification is based on STEM employment data between 2005 and 2014 while the BDS-HT statistics span 1977 to 2014. This means that BDS-HT statistics capture the current and historical performance of industries considered High Tech in the late 2000s. They do not capture the dynamics of industries that were High Tech, according to their employment of STEM workers, in the 1970s or 1980s. Since the method relies on the industries with the highest share of STEM workers, the set of industries identified are stable even over the 10 year window. Second, to overcome issues associated with changing industrial classification systems, the High Tech statistics leverage the work of Fort and Klimek (2016) (henceforth FK) to create longitudinally consistent industry classifications for establishments in the LBD. Moreover, while the FK codes allow industry codes to vary over time, GM impose conditions to select a single, time-invariant industry code for each establishment. This prevents industry switching at the establishment level.

Ultimately, High Tech industries account for about 4% of firms and 6% of employment. Figure 8 shows the evolution of the High Tech firm and employment shares over time. An important feature of the BDS data is that it provides employment, establishment, and firm flows by firm characteristics. Since industry classification is defined at the establishment level, multi-unit firms can be comprised of establishments operating in multiple industries. This implies that firms may be classified as both High Tech and non-High Tech. The firm share shown in 8 captures the share of firms with at least one establishment in a High Tech industry. Since the late 1970s the share of firms operating in High Tech industries has more than doubled from 2% to over 4%. The employment share, on the other hand, has remained relatively stable hovering around 6%.

Figure 8: High Tech Firm and Employment Share

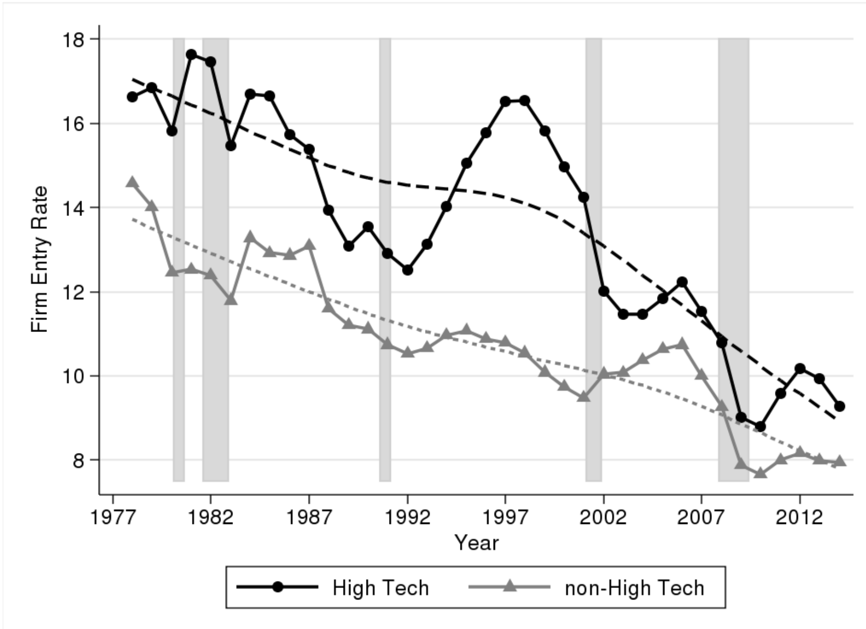


Source: Goldschlag and Miranda (2016)
Notes: Firm share calculated as the total number of firms with at least one High Tech establishment divided by the total number of firms in the economy. Employment share is total employment in High Tech establishments divided by total employment. Y-axis does not start at zero. Shaded regions represent recessions as defined by NBER.

One of the most striking features of the BDS High Tech data is the significant increase in entry and young firm activity through the 1990s, a period marked by strong productivity growth. Figure 9 shows the startup rate for the High Tech and non-High Tech sector. The startup rate in both the non-High Tech and High Tech sectors was declining in the mid to late 1980s. This is consistent with the literature focused on the secular decline in measures of business dynamism.[17] Starting in 1992, however, the startup rate in High Tech surges, peaking in 1998 before collapsing through the early 2000s. After 2002, the startup rate in High Tech starts to converge with the non-High Tech sector.

[17] See Decker, Haltiwanger, Jarmin, and Miranda (2014) for a detailed description of this literature.

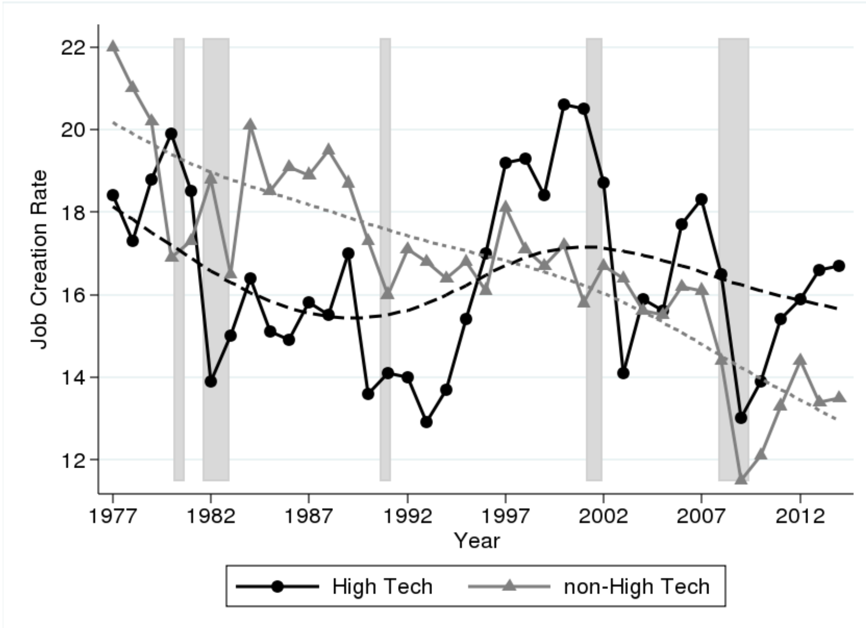
Figure 9: High Tech vs non-High Tech Firm Entry



Source: Goldschlag and Miranda (2016)
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.

The pattern of increased High Tech activity in the 1990s is also reflected in job creation rates. Figure 10 shows the job creation rate for the High Tech and non-High Tech sectors. Starting in the early 1990s the job creation rate in High Tech sector increases rapidly, remaining above that for the non-High Tech sector through 2002. After 2002, the job creation rate for the High Tech sector begins to resemble the job creation rate in the non-High Tech sector with the gap between the two growing larger at the end of the time series.

Figure 10: High Tech vs non-High Tech Job Creation Rate



Source: Goldschlag and Miranda (2016)
Notes: Trends calculated by applying a Hodrick-Prescott filter with a multiplier of 400. Job creation rate is total positive employment changes divided by the average employment in t-1 and t. Y-axis does not start at zero. Shaded regions represent recessions as defined by NBER.

The BDS High Tech data products provide an example of extensions to the BDS data infrastructure that focus specifically on an innovative population of firms. Businesses in the High Tech sector play an important role in reallocation by introducing new products and services that often have economy-wide impacts. The BDS High Tech data provide employment, establishment, and firm flows in the High Tech sector by firm characteristics such as firm age and firm size. With these data we can begin to paint a picture of the dynamics of businesses in the High Tech sector and therefore better understand the impact of innovative firms in the economy.

3 Additional Extensions

In addition to the projects described in Section 2, there are a number of extensions that could be made to capture additional dimensions of innovative activity in the economy. The underlying data infrastructure is flexible enough to accommodate any number of firm and establishment-level linkages. Here we highlight several projects on the horizon for the BDS-IF project.

First, copyright registration represents a natural counterpart to patents and trademarks. The economic importance of copyrights as a form of intellectual property has been studied in the literature, but relatively little has been done to harness copyrights as a measure of innovative activity (Landes and Posner, 1989; Laforet, 2008). Copyrights are not only used to protect creative content, such as novels or films, but also computer-code, maps, and recipes.[18] For instance, a 2014 ruling clarified that Application Programming Interfaces (APIs) can be protected by copyright (Oracle America, Inc. v. Google, Inc.). All of the largest ten companies by market capitalization have registered copyrights on things such as user manuals, shareholder statements, and branding designs. While creative companies clearly dominate this field, one might expect a manufacturing company to have copyright registration activity around the introduction of a new product. The text of marketing materials, text displayed on packaging, and product descriptions are all types of content that can be protected by copyright. In addition, if the product itself involves text or some kind of non-functional design, this too can be copyrighted. Thus, copyrights might offer an alternative measure of innovative activity in the economy–capturing certain types of activity not otherwise covered by patenting and trademarking.

Second, existing links between the LBD and the Survey of Industrial Research and Development (SIRD), the Business Research and Development and Innovation Survey (BRDIS), and in the future the BRDIS-M could be used to develop firm, establishment, and employment flows associated with R&D performing firms. Expenditures on R&D activities is one of the most well established measures of innovation in the literature, making it the logical next step for the BDS-IF project (Cohen, 2010). Foster, Grim, and Zolas (2016) find that R&D preforming firms are both larger and younger than the population as a whole, in contrast to

[18]Under current US law, copyrights for work created after 1978 owned by corporations last for 95 years from the date of first publication.

patenting firms which are both larger and older. Though the surveys are based on a sample of firms that changes over time, this result is consistent from 1995 to 2010. The authors also find that firms that engage in trade are much more likely to preform R&D, in keeping with trade literature that identifies a positive link between the size of the market, trade, R&D expenditures, and within-plant productivity growth (Aw, Roberts, and Xu, 2011).

Firms with more structured management practices also tend to be more productive (Bloom, Eifert, Mahajan, McKenzie, and Roberts, 2013). The Management and Organizational Practices Survey (MOPS) was designed to measure the extent to which firms employ structured management practices. MOPS covers about 37,000 manufacturing establishments and is by far the largest survey exclusively focused on management practices (Buffington, Foster, Jarmin, and Ohlmacher, 2017). Thus, it provides new opportunities to understand the dynamics of firms with different styles of management. Bloom, Brynjolfsson, Foster, Jarmin, Patnaik, Saporta Eksten, and Van Reenen (2013) find a link between more structured management and higher rates of innovation investment in IT. It seems natural then to use measures of management practices to capture an additional dimension of innovative activity in the economy.

4 Conclusion

The reallocation of labor and capital is an important driver of economic growth and rising productivity. Innovative firms, those that introduce new products, services, and/or business methods, contribute disproportionately to those reallocation flows. Innovative activity is notoriously difficult to measure, in large part because inputs and outputs of innovative activity are often novel, complex, and difficult to observe. In this paper we describe the BDS-IF project, a measurement agenda at the Census Bureau focused on the business dynamics of innovative firms. Motivated by the many different dimensions of innovative activity in the economy, we take an expansive approach focusing on both inputs (e.g., R&D spending, hiring STEM workers) and outputs (e.g. patents, trademarks) of the innovation process. We describe three active projects extending the BDS infrastructure to measure the dynamics of patenting firms, trademarking firms, and firms in High Tech industries. We outline a number of interesting findings gleaned from each and describe directions for future work including the use copyright data, R&D data, and information on management practices. Ultimately, these efforts will produce public use statistics on innovative firms that will be valuable to both researchers and policy makers.

References

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

Acs, Z. J., and D. B. Audretsch (1991) "R&D, firm size and innovative activity," *Innovation and technological change: An international comparison*, 98(2), 451–456.

Aghion, P., and P. Howitt (1992) "A model of growth through creative destruction," *Econometrica*, 60(2), 323–351.

Akcigit, U., and R. W. Kerr (2017) "Growth through Heterogeneous Innovations," *Journal of Political Economy*.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2011) "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, 101(4), 1312–44.

Balasubramanian, N., and J. Sivadasan (2010) "NBER Patent Data-BR Bridge: User guide and technical documentation," .

Bloom, N., E. Brynjolfsson, L. Foster, R. S. Jarmin, M. Patnaik, I. Saporta Eksten, and J. Van Reenen (2013) "IT and Management in America," *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-01*.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013) "Does management matter? Evidence from India," *The Quarterly Journal of Economics*, 128(1), 1–51.

Buffington, C., L. Foster, R. Jarmin, and S. Ohlmacher (2017) "The management and organizational practices survey (MOPS): An overview 1," *Journal of Economic and Social Measurement*, 42(1), 1–26.

Cohen, W. M. (2010) "Fifty years of empirical studies of innovative activity and performance," *Handbook of the Economics of Innovation*, 1, 129–213.

Cummins, C., B. H. Hall, E. S. Laderman, and J. Mundy (1985) "The R&D master file: Documentation," *Manuscript. Cambridge, Mass.: NBER*.

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., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2017) "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown," .

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016) "Where has all the skewness gone? The decline in high-growth (young) firms in the US," *European Economic Review*, 86, 4–23.

Dinlersoz, E., N. Goldschlag, A. Myers, and N. Zolas (2017) "An Anatomy of Trademarking by Firms in the United States," in *CRIW: Measuring and Accounting for Innovation in the 21st Century*, ed. by C. Corrado, J. Miranda, J. Haskel, and D. Sichel. NBER.

Fort, T. C., and S. D. Klimek (2016) "The Effect of Industry Classification Changes on US Employment Composition," Discussion paper.

Foster, L., C. Grim, J. Haltiwanger, and Z. Wolf (2017) "Innovation, Productivity Growth and Productivity Dispersion," in *CRIW: Measuring and Accounting for Innovation in the 21st Century*, ed. by C. Corrado, J. Miranda, J. Haskel, and D. Sichel. NBER.

Foster, L., C. Grim, and N. J. Zolas (2016) "A Portrait of Firms that Invest in R&D," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-41*.

Foster, L., and P. Norman (2016) "The Annual Survey of Entrepreneurs: An Introduction,"

US Census Bureau Center for Economic Studies Paper No. CES-WP-15-40R.
Goldschlag, N., and J. Miranda (2016) "Business Dynamics Statistics of High Tech Industries," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-55*.
Graham, S. J., C. Grim, T. Islam, A. C. Marco, and J. Miranda (2015) "Business dynamics of innovating firms: Linking US patents with administrative data on workers and firms," .
Graham, S. J., G. Hancock, A. C. Marco, and A. F. Myers (2013) "The USPTO trademark case files dataset: Descriptions, lessons, and insights," *Journal of Economics & Management Strategy*, 22(4), 669–705.
Griliches, Z. (1998) "Patent statistics as economic indicators: a survey," in *R&D and productivity: the econometric evidence*, pp. 287–343. University of Chicago Press.
Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001) "The NBER patent citation data file: Lessons, insights and methodological tools," Discussion paper, National Bureau of Economic Research.
Hecker, D. E. (2005) "High-Technology Employment: a NAICS-Based Update," *Monthly Lab. Rev.*, 128, 57.
Jarmin, R. S., and J. Miranda (2002) "The longitudinal business database," .
Kerr, W. R., and S. Fu (2008) "The survey of industrial R&Dpatent database link project," *The Journal of Technology Transfer*, 33(2), 173–186.
Laforet, S. (2008) "Size, strategic, and market orientation affects on innovation," *Journal of business Research*, 61(7), 753–764.
Landes, W. M., and R. A. Posner (1989) "An economic analysis of copyright law," *The Journal of Legal Studies*, 18(2), 325–363.
Locke, R. M., and R. L. Wellhausen (2014) *Production in the innovation economy*. MIT Press.
OECD (2007) "Science, Technology and Innovation Indicators in a Changing World: Responding to Policy Needs," Discussion paper, OECD.
——— (2010) "Measuring Innovation: A New Perspective," Discussion paper, OECD.
Romer, P. M. (1990) "Endogenous technological change," *Journal of political Economy*, 98(5, Part 2), S71–S102.
Scherer, F. M. (1965) "Firm size, market structure, opportunity, and the output of patented inventions," *The American Economic Review*, 55(5), 1097–1125.
Schumpeter, J. A. (1942) *Capitalism, socialism and democracy*. New York: Haper & Brothers.
Smith, K. (2005) "Measuring innovation," .
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.
Thompson, P., and M. Fox-Kean (2005) "Patent citations and the geography of knowledge spillovers: A reassessment," *American Economic Review*, pp. 450–460.
Zolas, et al. (2015) "Wrapping it Up in a Person: Examining Employment and Earnings Outcomes for Ph.D. Recipients," *Science*, 350(6266), 1367–1371.