University of Toronto
Department of Economics

Working Paper 686

The Role of Nonemployers in Business Dynamism and Aggregate Productivity

By Pedro Bento and Diego Restuccia

January 19, 2021

The Role of Nonemployers in Business Dynamism and Aggregate Productivity[†]

Pedro Bento
Texas A&M University[*]

Diego Restuccia
University of Toronto
and NBER[**]

January 2021

Abstract

The well-documented decline in business dynamism, measured in the literature by the net entry rate of employer firms, has been proposed as an explanation for the productivity growth slowdown in the United States. We assess the role of nonemployers, firms without paid employees, in business dynamism and aggregate productivity. Including nonemployers, the total number of firms has instead increased since the early 1980s, which in the context of a standard model of firm dynamics implies an average annual growth of aggregate productivity of 0.26%, one-quarter of the productivity growth in the data. Accounting for changes in the share of nonemployers and the firm size distribution over time, the increase in the total number of businesses implies an even higher productivity growth of 0.52% annually. The productivity growth slowdown is not due to changes in business dynamism.

Keywords: nonemployers, employer firms, business dynamism, productivity, TFP.
JEL codes: O4, O51, E1.

[†]For helpful comments and suggestions we thank Benjamin Bridgman, John Haltiwanger, Pete Klenow, Benjamin Pugsley, and audiences at Texas A&M University, the University of Texas at Austin, the 2019 Midwest Macroeconomic Meetings at Michigan State University, the 2019 RCEA Growth, Innovation and Entrepreneurship Conference at Wilfrid Laurier University, and the 2020 Econometric Society North American Winter Meetings in San Diego. All errors are our own. Restuccia gratefully acknowledges the support from the Canada Research Chairs program and the Bank of Canada Fellowship. The views expressed herein are not necessarily those of the Bank of Canada and are the authors' alone.

[*]Department of Economics, Texas A&M University, 254 Liberal Arts Social Sciences Building, 4228 TAMU, College Station, TX 77843, USA. E-mail: pbento@tamu.edu.

[**]Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada. E-mail: diego.restuccia@utoronto.ca.

1 Introduction

A number of studies have documented a slowdown in business startups and entrepreneurship in the United States over the last several decades. The decline since the Great Recession, in particular, has been proposed as a potential source of the growth slowdown in aggregate productivity (Decker et al., 2016; Furman and Orszag, 2018). However, Decker et al. (2016) and Li (2017) have noted that standard measures of business dynamism appear unrelated to estimates of aggregate total factor productivity (TFP) growth before the Great Recession. In this paper, we construct a broader measure of the total number of firms that provides a more comprehensive picture of business dynamism in the United States. We use this measure to assess the quantitative contribution of net firm entry on aggregate productivity growth.

Canonical theories of firm size and firm dynamics, such as Hopenhayn (1992), are used to draw implications for aggregate TFP from data on business dynamism. In these models, aggregate output depends on aggregate factor inputs and a term that aggregates the productivity of all firms, which also depends on the total number of firms. In this context, measured TFP, the aggregate amount of output per unit of composite aggregate inputs, depends on the total number of firms. We construct a comprehensive measure of U.S. businesses that includes nonemployers, that is businesses that are subject to federal income tax but have no paid employees, composed solely of owner-managers and unpaid workers such as family members. We show that this measure of the number of firms has increased substantially. Typical measures of business dynamism are based on employer firms, firms with at least one paid employee. But nonemployers account for 82% of all firms in 2014, suggesting their evolution over time is an important determinant of changes in the total number of firms.

We combine employers data from the Business Dynamics Statistics (BDS) with nonemployers data from the Nonemployer Statistics (NES) and other data sources, to construct a measure of the total number of U.S. businesses from 1981 to 2014. We focus on the number of firms per worker, which in theory is the relevant measure when drawing implications for aggregate productivity (Hopenhayn, 1992; Karahan et al., 2019). Although the number of employers per worker decreased by 4.5% from 1981 to 2014, consistent with the findings in Karahan et al. (2019), the total number of firms per worker in our measure increased by 53% over the same period.

We consider a standard model of firm dynamics with endogenous entry in order to quantitatively assess the impact of the surge in the number of firms on aggregate TFP. In the baseline model, the distribution of firm-level productivities is constant over time but we allow for changes in aggregate employment as observed in the data. We begin by assuming that the cost of entry changes over time (along with aggregate employment) in order to match the observed evolution in the number of firms; and study the aggregate productivity implications. The calibrated model implies that aggregate TFP grew at an annualized rate of 0.26% from 1981 to 2014 due to the increase in the number of firms, one quarter of the actual growth in measured TFP. In contrast, using the number of employers as is standard in the literature, the model implies slightly negative growth in TFP. Over 33 years, the implied rates compound into a cumulative increase in aggregate TFP of 9% when using the total number of firms per worker and −1% when using the number of employers per worker.

The baseline model understates the growing share of nonemployers. It also implies a time invariant employment size distribution, while the data and recent studies of market concentration suggest otherwise (Autor et al., 2020; Rossi-Hansberg et al., forthcoming). To accommodate these changes in the data, we extend the baseline model to allow for changes in average productivity over time arising from differential exit rates between employers and nonemployers, and from changes over time in the exit and productivity growth rates of incumbents. We discipline the model using data on differential exit rates between employers and nonemployers, and data on changes in the employment-size distribution of entrants and incumbents, as well as changes in the revenue share of nonemployers. Under relatively weak structural assumptions commonly made in the firm-dynamics literature, we show how this additional data can be combined with data on the total number of firms to derive the implied change in the average productivity of firms over time up to a constant, that is, relative to any change in productivity common to all firms. We find that the implied average firm-level productivity has increased from the 1980s to 2014, making the cumulative increase in aggregate TFP in the extended model 17.9% compared to 9% in the baseline model. We also show that TFP growth implied by the extended model correlates well with observed TFP growth over the medium and longer run relative to a model that includes only employer firms.

Our comprehensive measure of business dynamism that includes nonemployers complements the important work of Decker et al. (2014) and Decker et al. (2016), who emphasize the decline in net entry rates for employer firms. We show that the net entry rate of all firms has not declined. Karahan et al. (2019) and Hopenhayn et al. (2019) document a marginally declining trend in employers per worker and conclude that changing business dynamism has not been a quantitatively important driver of TFP trends. We show instead that when incorporating nonemployers in the total number of firms, business dynamism has contributed substantially to productivity growth. Karahan et al. (2019) and Hopenhayn et al. (2019) also document an increase in the average age of employer firms over time, driven by lower exit rates. We show that accounting for lower exit rates over time increases the implied contribution of business dynamism to aggregate TFP growth, as older firms tend to be much more productive than young firms. Pugsley and Şahin (2019) provide evidence that growth in the numbers of employer and nonemployer firms tend to move together over time between 1997 and 2012, and that these growth rates are correlated with aggregate employment. We extend their analysis by looking at nonemployers over a longer time period, and show that while annual growth rates are correlated, the number of nonemployers grew much faster than the number of employers.

The literature has also documented a drop over time in the job reallocation rate of employer firms, measured by the sum of job creation and job destruction normalized by aggregate employment, as another metric of the decline in business dynamism (Decker et al., 2014, 2016). Hyatt et al. (2020) construct a new measure of job reallocation that accounts for flows to and from nonemployers in addition to employers, and find the decline over time in the job reallocation rate is smaller compared to measures that abstract from nonemployers. We instead focus on the aggregate productivity impact of changes in business dynamism arising from changes in the total number of firms per worker, the share of nonemployers, and the employment size distribution. A recent literature has identified policy and institutional distortions in developing countries that encourage more firm entry while distorting the allocation of labor across firms and thereby lowering aggregate TFP (Hsieh and Klenow, 2014; Bento and Restuccia, 2017, forthcoming; Bento, 2020). We assess whether changes in distortions may be driving trends in business dynamism in the United States using establishment data on employment and revenue. We argue that changes in distortions are not important in accounting for the U.S. experience in recent decades.

In the next section, we discuss the evidence on nonemployers and why they may matter for business dynamism. Section 3 describes the data for employers and nonemployers and documents trends in the variables of interest. In Section 4, we describe our baseline model of firm entry to assess the quantitative impact of firms per worker on aggregate productivity. Section 5 extends the analysis to include differential firm exit rates and firm-level productivity growth showing that implied TFP growth from firm dynamics is even larger than in the baseline model. Section 6 provides a discussion of potential alternative drivers of business dynamism. We conclude in section 7.

2 Nonemployer U.S. Businesses

We construct a comprehensive measure of the total number of firms in the U.S. economy to assess the role of changes in net firm entry on aggregate productivity. We focus on a measure of the total number of firms that includes nonemployer businesses. Nonemployer businesses are firms with no paid employees, including self-employed entrepreneurs. A comprehensive measure of firms may be relevant in understanding changes in net entry rates over time, as is the case when considering very small firms in the context of cross-country differences in establishment size (Bento and Restuccia, 2017, forthcoming).

For many economic questions it is reasonable to abstract from nonemployers, as they contribute little to aggregate output in the U.S. economy. Although nonemployers constitute 82% of all U.S. businesses in 2012, they represent only about 3% of total revenues. However, theories of firm size and firm dynamics suggest patterns of firm entry and exit are essential for aggregate productivity implications. In this context, it is important for the analysis to account for all firms. This is the case even if nonemployers are less productive than employer firms and account for a small proportion of output and employment, although these characteristics need to be taken into account.

Including nonemployers in the total measure of firms raises important questions. Are nonemployer firms using different technologies than employer firms or operating in different product markets? Or are nonemployers the same as employer firms albeit with lower productivity? Our data, together with recent papers by Acs et al. (2009), Davis et al. (2009), and Fairlie et al. (2018), provide a characterization of U.S. nonemployers that can be compared with employer firms. In the data, nonemployers coexist and compete with employers within narrow industries. In each of the nine industries we consider, nonemployers represent more than 55% of all firms in 2007 (more than 80% in five industries) and are more prevalent in industries with smaller average employment firm size. The survival rate of nonemployer startups is close to that of small employer startups. Data on employment in nonemployers (i.e., owner-managers and unpaid workers) are not available, but average growth rates of revenue are similar to that of small employer firms. A small percentage of nonemployers transition into employer status each year, roughly consistent with employment growth rates among small employers. Nonemployers compare to small employers in similar proportions than small employers compare to medium or large employers. For instance, in terms of average revenue per firm in 2007, nonemployers are about 12% of that of small employers (less than 5 employees). Small employers in turn are about 17% of that of medium size employers (10 to 19 employees), which in turn are 16% of that of large employers (50 to 99 employees). Nonemployers certainly don't resemble large employers in terms of average revenue, but the difference between nonemployers and small employers resembles the difference between small and large employers. The main difference between employer and nonemployer firms other than size appears to be their rate of exit, consistent with a decreasing exit rate with size (Haltiwanger et al., 2013). In particular, while about 8 to 9% of all employers exit each year, the exit rate of all nonemployers is higher, about 15% (Davis et al., 2009).

From these facts we conclude it is reasonable to treat nonemployer businesses as employers but operating at a lower scale, possibly because of lower productivity, with higher exit rates.

3 Data

We describe the data and procedure used to construct our measure of the total number of firms over time in the U.S. economy. Data for employer firms is from the U.S. Census Bureau's Business Dynamics Statistics (BDS), the standard data source in the business dynamism literature (Decker et al., 2014). The employer data contains employer-firm counts by industry, employment size, and age from 1977 to 2014. All non-farm firms with at least one formal employee are included.

Data for nonemployers is from the U.S. Census Bureau's Nonemployer Statistics (NES). NES contains economic data for businesses that have no paid employees and are subject to federal income tax, providing nonemployer business counts by industry for 1992 and from 1997 onward. The U.S. Internal Revenue Service (IRS) tax return data is used by the Census Bureau to identify the universe of potential nonemployers. IRS counts up to 2008 are reported in U.S. Statistical Abstracts. Care is then taken to identify duplicates (multiple tax numbers belonging to one firm), and reclassify nonemployers when they are properly part of an employer firm.

To construct our measure of the total number of firms, we simply add nonemployer businesses to employer firms. This is done for the years 1992 and 1997 to 2014 for which we have data for both nonemployers and employer firms. We impute nonemployer counts for the years 1981 to 1991 and 1993 to 1996. For the years 1993 to 1996, we simply assume that the number of nonemployers increased smoothly from 1992 to 1997, and add the implied number of nonemployers to the observed number of employers. For the years 1981 to 1991, we impute the number of nonemployers using IRS data. We describe this imputation in Appendix A.

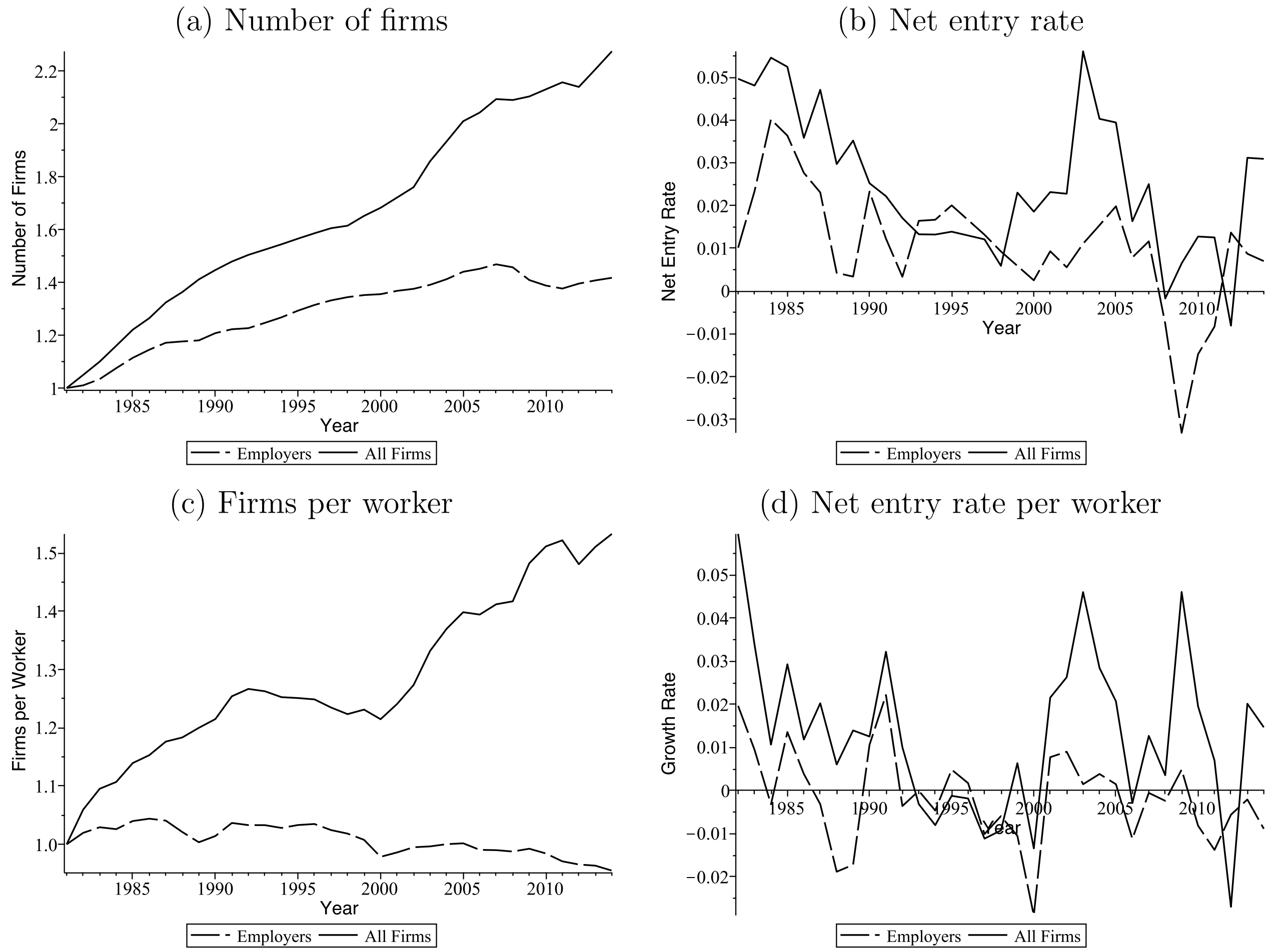
Figure 1 documents the evolution of the number of firms and firms per worker in the United States. Panel (a) reports our measure of the number of firms and the more common measure of the number of employers over time, normalized to one in 1981. Panel (b) reports the net entry rate (growth in the number of firms) of all firms and employers. Two features of the data stand out. First, the net entry rate of all firms has been consistently higher than that of employer firms. Second, the net entry rate of all firms declined along with that of employer firms from the early 1980s, but then diverged sharply starting in the late 1990s. From 1981 to 2014, while the number of employers increased by 42%, the total number of firms increased by a striking 128%.

Theories of firm dynamics suggest the more relevant measure of business dynamism when drawing implications for TFP is the number of firms per worker. Using data on the total employed civilian non-institutional population from the U.S. Bureau of Labor Statistics' Current Population Survey (CPS), Figure 1 panel (c) documents the number of firms per worker for all firms and for employers, and panel (d) the net entry rate per worker. Both the total number of firms per worker and the number of employers per worker drop during the 1990s. But whereas the growth rate of employer firms per worker stays negative (on average) after 2000, the total number of firms per worker recovers and grows at a positive rate. From 1981 to 2014, the number of firms per worker increases by 53%, whereas the number of employers per worker decreases by 4.5%.[1]

The striking difference in the number of firms over time between all firms and employers is robust to removing sole-proprietors from nonemployer counts. Data on the legal form of nonemployers is available from 1997 onwards. Figure 2 compares the number of all firms per worker, the

[1]The measure of CPS employment we use is fairly consistent with the measures of labor force participation analyzed in Karahan et al. (2019) and Hopenhayn et al. (2019).

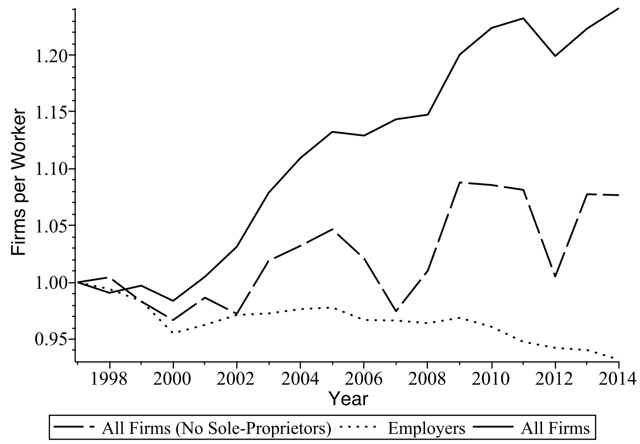
Figure 1: Evolution of Firms and Firms per Worker



Notes: Panel (a) reports the number of all firms and the number of employer firms with levels normalized to one in 1981. Panel (b) reports the net entry rate of all firms and of employer firms. Panel (c) reports the number of firms per worker and panel (d) the net entry rate per worker.

number of firms without nonemployer sole-proprietors, and the number of employers. Although the cumulative increase in the number of firms per worker since 1997 is lower when removing sole-proprietors, an 8% increase rather than a 24% increase with all firms, it is still markedly higher than the −7% for employers per worker. We also note the increase in the number of firms per worker over time occurs within sectors and is not the result of reallocation across sectors with different firms per worker. We document in Appendix B that firms per worker have increased in seven out of nine sectors and that most employment reallocation has occurred between manufacturing and other services with similar increases in firms per worker. More formally, we show via a counterfactual that only 20% of the increase in the number of firms per

Figure 2: Number of Firms per Worker with and without Sole-Proprietors



Notes: The solid line represents the total number of firms per worker in the data, the doted line is employer firms per worker, and the dashed line is the total number of firms per worker excluding sole-proprietors. In each case, levels are normalized to 1 in 1997.

worker can be accounted for by the change in economic structure during the period.

In summary, accounting for nonemployer businesses in firm counts dramatically changes the pattern of net entry over time in the U.S. economy. While the number of employers per worker has fallen slightly over the last three decades, the number of all firms per worker has risen substantially by more than 50%.

4 Baseline Model

We consider a version of the firm dynamics model in Hopenhayn (1992) in order to provide a mapping from changes in net entry of firms to aggregate TFP. We also use the model to assess the factors leading to the divergence over time between the number of employer and nonemployer firms.

4.1 Environment

At each date, a single homogeneous good (the numéraire) is produced by firms. Firms have access to a decreasing returns to scale technology in variable inputs and are heterogeneous with respect to their productivity z:

$$y = (Az)^{1-\alpha}\ell^{\alpha}, \tag{1}$$

where y is output, ℓ is the labor input, and A an exogenous productivity term common to all firms that can change over time. Decreasing returns to scale in variable inputs implies $\alpha \in (0, 1)$, hence the optimal scale of a firm depends non-trivially on productivity. More productive firms operate at a larger scale by hiring more inputs, producing more output, and generating higher profits. Firms take the current real wage w as given, and the only cost incurred by incumbents is their wage bill.

There are a large number of potential entrants that can become producers by incurring an entry cost equal to $c_e \cdot Y/L$, where Y/L is aggregate output per worker.[2] We allow c_e to change over time in order to match the evolution of the net entry rate of firms as we explain below. We assume potential entrants draw their productivity z from some constant cumulative distribution function $G(z)$, and learn their productivity after entry. There is no fixed operating cost for producers, and firm-level productivity is assumed to stay fixed over the lifetime of a firm. We assume all firms face an exogenous probability of exit λ in each period after production.

We denote employment by L which can change over time. We assume that firms always believe current levels of L, c_e, and A persist forever. At the beginning of each period they learn otherwise, but then again believe current levels of these variables persist. We discuss below and elaborate in Appendix C that this assumption about beliefs does not affect our main results other than the implied value of entry costs. We abstract from household choices by assuming a constant exogenous real interest rate R.

4.2 Equilibrium

A *competitive equilibrium* is defined by a wage rate w, firm-level functions labor demand $\ell(z)$ and per-period profits $\pi(z)$, and number of firms N, given exogenous entry cost $c_e \cdot Y/L$, labor

[2] Having entry costs scale up with output per capita is consistent with the evidence in Bollard et al. (2016) and Bento and Restuccia (forthcoming).

supply L, real interest rate R, and firm-level productivity distribution $G(z)$, such that:

(i) Given w, firms choose $\ell(z)$ to maximize $\pi(z)$.

(ii) Free entry ensures the expected present value of lifetime profits for an entrant is equal to the entry cost, $\int_z \frac{\pi(z)}{1-\rho} dG(z) = c_e \cdot Y/L$, where $\rho = (1-\lambda)/(1+R)$.

(iii) The labor market clears: aggregate labor is equal to the quantity of labor demanded by firms, $L = N \int_z \ell(z) dG(z)$.

The equilibrium in each period can be easily solved. Producers choose labor to maximize operating profits, resulting in the following optimal demand for labor, output, and operating profits, expressed as functions of z;

$$\ell = Az \left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}}, \tag{2}$$

$$y = Az \left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}, \tag{3}$$

$$\pi = Az \left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}} (1-\alpha). \tag{4}$$

Labor market clearing implies total labor is equal to aggregate labor demand;

$$L = N \cdot A \left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}} \bar{z},$$

where $\bar{z}$ is equal to average firm-level productivity. This is in turn equal to the expected value of each draw z from $G(z)$, or $\int_z z \, dG(z)$. The wage can therefore be expressed as a function of the number of firms per worker (N/L);

$$w = \alpha (A\bar{z})^{1-\alpha} \left(\frac{N}{L}\right)^{1-\alpha}. \tag{5}$$

Using equations (3) and (5), aggregate output per worker as a function of firms per worker is;

$$\frac{Y}{L} = \frac{N}{L} \cdot A \left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}} \bar{z} = (A\bar{z})^{1-\alpha} \left(\frac{N}{L}\right)^{1-\alpha}. \tag{6}$$

Free entry ensures N in each period is such that the discounted expected profits of an entrant are equal to the cost of entry:

$$c_e \cdot \frac{Y}{L} = \frac{A(1-\alpha)\bar{z}}{1-\rho} \left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}. \tag{7}$$

This free entry condition holds as long as c_e is not too high relative to previous periods. If c_e were to increase too much from period $t-1$ to period t, then the number of firms in period t would be $N_t = (1-\lambda)N_{t-1}$. This constraint never binds in practice in our application. Along with equations (5) and (6), the free entry condition implies the following characterization of the number of firms per worker;

$$\frac{N}{L} = \frac{1-\alpha}{c_e(1-\rho)}. \tag{8}$$

Note that firms per worker does not depend on the common productivity term A, on average firm-level productivity $\bar{z}$, or on the size of the workforce L.

Lastly, output per capita, aggregate TFP, and the wage can be expressed as functions of exogenous variables;

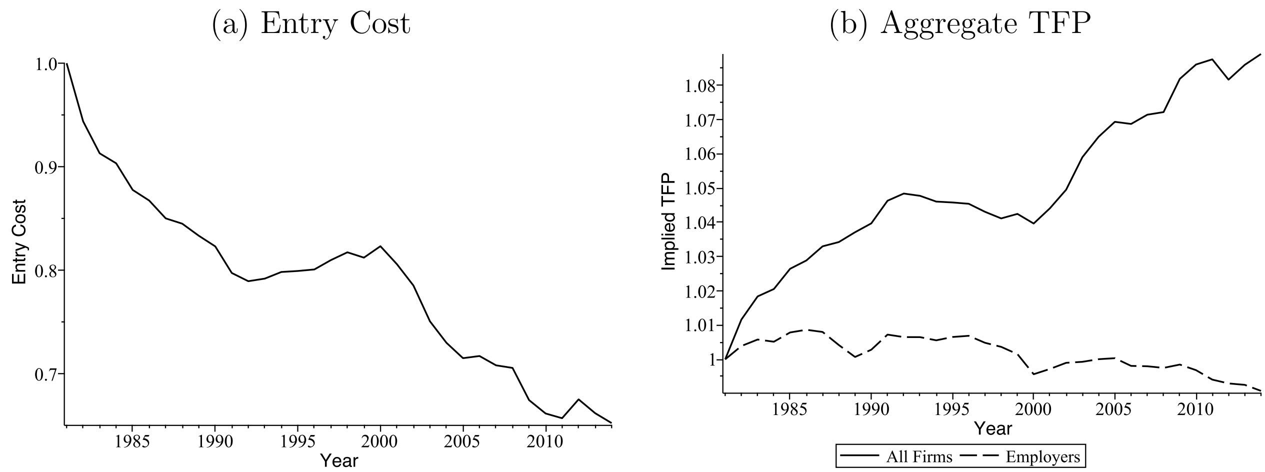
$$\frac{Y}{L} = TFP = (A\bar{z})^{1-\alpha} \left(\frac{1-\alpha}{c_e(1-\rho)}\right)^{1-\alpha}, \tag{9}$$

$$w = \alpha \cdot TFP. \tag{10}$$

4.3 Implications

Entry costs. In the model, the number of firms depends on employment L and entry cost c_e, whereas the number of firms per worker depends only on entry costs as indicated in equation (8). Hence, given the employment data, we infer changes in entry costs to match changes in the

Figure 3: Implied Entry Costs and Aggregate TFP in Baseline Model



Notes: Entry costs c_e calibrated to match the changes in total number of firms per worker in the data, normalized to 1 in 1981. Aggregate TFP in each year relative to 1981 is from equation (11) using the total number of firms per worker (solid line) and the number of employer firms per worker (dashed line).

number of firms per worker in the data and normalizing the initial entry cost to one. Figure 3a documents the implied evolution of c_e which follows the inverse of the number of firms per worker. We note that the implied entry cost depends on assumptions about the beliefs held by firms with respect to future aggregate employment growth and future entry costs. But regardless of what we assume about firms' expectations, the overall decline in the implied entry cost still holds. We show in Appendix C how entry costs change when firms know the future paths of c_e and L.

Aggregate productivity. Aggregate labor productivity and aggregate TFP coincide in the model and are characterized by equation (9) as a function of parameters and exogenous variables such as the entry cost, or by equation (6) as a function of exogenous variables and the number of firms per worker. As a result, we connect changes in the number of firms per worker in the data to changes in aggregate TFP, noting that average productivity $\bar{z}$ drops out since we assume a constant distribution of productivity:

$$\frac{TFP_t}{TFP_{1981}} = \left(\frac{A_t}{A_{1981}}\right)^{1-\alpha} \left(\frac{N_t/L_t}{N_{1981}/L_{1981}}\right)^{1-\alpha}. \tag{11}$$

To calculate the implied aggregate TFP over time (relative to 1981), we use a value for α equal to 0.8, consistent with much of the firm-dynamics literature.[3] Figure 3b illustrates the main result on the implied TFP in equation (11) using the total number of firms per worker (solid line), contrasting it with the the number of employers per worker (dashed line) as an alternative measure of the number of firms. The increase in the total number of firms per worker from 1981 to 2014 implies a 9% cumulative increase in TFP, or an annualized growth rate of 0.26%. This is substantial relative to the observed 0.85% annual growth rate of TFP in the United States over the same time period.[4] In other words, over one quarter of the growth in TFP during the period can be attributed to the change in the number of firms per worker. In contrast, using the number of employers per worker as is standard in the literature, the implied TFP decreases by 1% between 1981 and 2014.

Evolution of nonemployer firms. To derive the implications of the model for the share of employer and nonemployer firms, we assume that nonemployers are firms with optimal labor demand less than one, with labor provided by an owner-manager. We derive the demand for labor by firms by combining the firm's optimal demand for labor (equation 2), which is increasing in productivity z and decreasing in the wage, with the market clearing wage rate (equation 5), which is increasing in the number of firms per worker and average firm-level productivity $\bar{z}$;

$$\ell = \left(\frac{z}{\bar{z}}\right)\left(\frac{N}{L}\right)^{-1}. \tag{12}$$

If the number of firms per worker increases, demand for labor increases and pushes up the equilibrium wage (equation 5). This results in lower employment for any given level of productivity. Using equation (12) we can solve for a productivity threshold z_1 below which a firm is counted as a nonemployer;

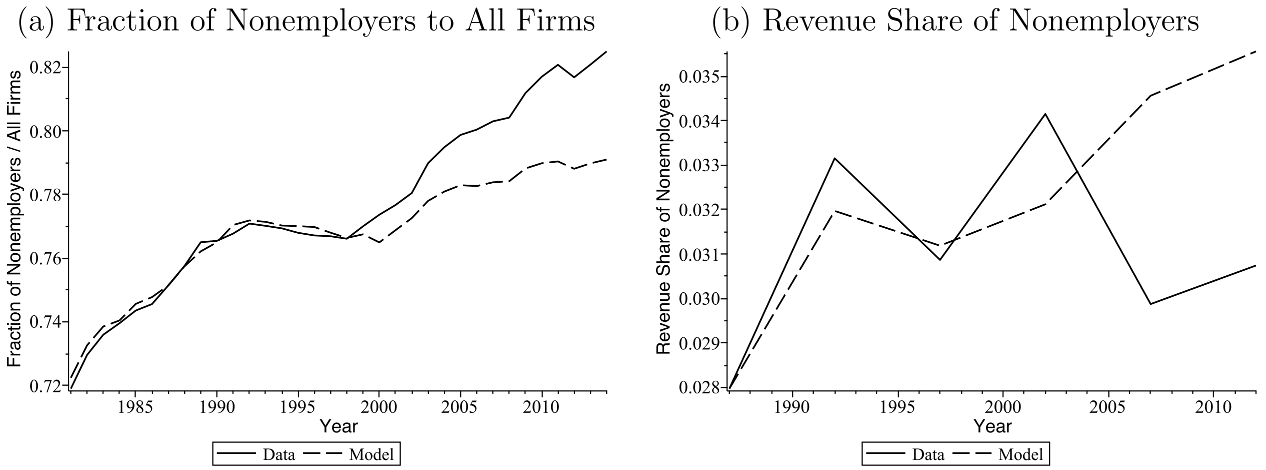
$$z_1 = \bar{z}\left(\frac{N}{L}\right). \tag{13}$$

The number of nonemployers per worker is the number of firms below the productivity threshold

[3] Note that $\alpha = 0.8$ implies a conservative estimate of TFP growth, as lower values occasionally in the literature imply larger effects on TFP growth from changes in the number of firms per worker.

[4] The measure of TFP growth in the data is utilization-adjusted TFP growth from Fernald (2012), updated to 2014.

Figure 4: Nonemployer Firms in the Model and Data



z_1, which is increasing in the number of firms per worker. If the number of firms per worker increases, the fraction of firms that are nonemployers $G(z_1)$ increases as z_1 rises. We calibrate $G(z)$ to U.S. data, targeting BDS data on the firm size distribution for employers and data on the revenue share of nonemployers in 1987. We provide more details in Appendix D.

Figure 4a compares the fraction of firms classified as nonemployers in the model to the data. The fraction of nonemployers in the model tracks the data closely until the late 1990's as the fraction of nonemployers grows from 72% in 1981 to 77% in the 1990s. After 1998 the fraction of nonemployers increases more in the data than in the model – the data shows an increase in the share of nonemployers to 82% by 2014, while the model implies a smaller increase to 79%. Figure 4b illustrates the same comparison between the model and the data for the revenue share of nonemployers. Note that revenue data for nonemployers and employers is only available during census years starting in 1987. The model implies a relatively steady increase in the revenue share of nonemployers from 2.8% to 3.6%. Although the model-implied share follows the data relatively closely up to 2002, it then increases more than the data by 2012.

To summarize, the increase in firms per worker in the data can account for much of the change over time in the share of nonemployer firms, but starting in the late 1990s other forces may be contributing to changes in the firm size distribution, for instance changes in exit rates or in the productivity distribution which are assumed constant in the model. We assess the importance of changes in the firm size distribution for productivity implications in the next section.

5 Extended Model

The baseline model does not capture the entire shift observed over time in the share of nonemployer firms, potentially affecting the productivity implications. In addition, the baseline model features a constant distribution of employment across firms, whereas recent studies highlight a shift in employment from small to large firms.[5] Indeed, in our data the average employment size of all incumbent firms relative to entrants rose from 4 in the early 1980s to more than 6 by 2014. The shift in employment shares may reflect an increase in the average productivity of incumbents relative to entrants.[6]

We extend the baseline model to account for observed changes in the employment size distribution and the share of firms that are nonemployers and reassess the productivity impact of changes in business dynamism. For nonemployers, we continue to assume that productivity does not change over their life cycle and average productivity is constant over time. We now assume that the exogenous and constant exit rate of nonemployers is higher than that of employer firms. For employers, we allow for exogenous but time-varying exit and productivity growth rates. From the BDS data we know that exit rates for employers have changed over time. Given these observed exit rates and the resulting observed employment distribution, it is possible to infer productivity growth rates for incumbent employer firms in each year. But in what follows, we show how observations on the number of entrants and incumbents, combined with aggregate employment shares of incumbents relative to entrants, are sufficient in our framework to infer changes in average productivity over time.

To derive implications for aggregate productivity over time, we use equation (6) with 1982 as

[5]This shift in employment across firms is consistent with a recent literature documenting broader trends in market concentration of economic activity at the top of the firm distribution (Autor et al., 2020; Rossi-Hansberg et al., forthcoming).

[6]For instance Hopenhayn et al. (2019) note that average exit rates for incumbents have fallen over time and Aghion et al. (2020) point to a falling cost of expansion for incumbents.

the benchmark year since we now rely on firm exit:

$$\frac{TFP_t}{TFP_{1982}} = \left(\frac{A_t}{A_{1982}}\right)^{1-\alpha} \left(\frac{\bar{z}_t}{\bar{z}_{1982}}\right)^{1-\alpha} \left(\frac{N_t/L_t}{N_{1982}/L_{1982}}\right)^{1-\alpha}. \tag{14}$$

Changes in TFP can be decomposed into: (a) changes in the common productivity term A, (b) changes in average productivity across firms, and (c) changes in the number of firms per worker. We continue to use our measure of the total number of firms per worker over time for N/L, but unlike in the baseline model where average firm-level productivity $\bar{z}$ was assumed constant, in the extended model we use additional data to infer the evolution of $\bar{z}$. We now describe the specifics of our approach.

The model provides a mapping from size to productivity (up to a constant) given by equation (12). Denoting the average employment size of entrants by $\bar{\ell}_{ent}$ and all firms by $\bar{\ell}$, average productivity of all firms relative to entrants in each year is;

$$\frac{\bar{z}_t}{\bar{z}_{ent,t}} = \frac{\bar{\ell}_t}{\bar{\ell}_{ent,t}} \qquad \Rightarrow \qquad \bar{z}_t = \bar{z}_{ent,t}\frac{\bar{\ell}_t}{\bar{\ell}_{ent,t}}.$$

Since entrants draw z from a constant distribution $G(z)$ (their average productivity is constant), we infer average productivity of all firms in each year relative to 1982 as;

$$\frac{\bar{z}_t}{\bar{z}_{1982}} = \frac{\bar{z}_{ent,t}}{\bar{z}_{ent,1982}} \cdot \left(\frac{\bar{\ell}_t/\bar{\ell}_{ent,t}}{\bar{\ell}_{1982}/\bar{\ell}_{ent,1982}}\right) = \frac{\bar{\ell}_t/\bar{\ell}_{ent,t}}{\bar{\ell}_{1982}/\bar{\ell}_{ent,1982}}. \tag{15}$$

Our measure of $\bar{\ell}_t$ is simply L_t/N_t that includes all employer and nonemployer firms, whereas $\bar{\ell}_{ent,t} = L_{ent,t}/N_{ent,t}$. The BDS data provide information for entrant employer firms, however, the NES data do not distinguish between entrant and incumbent nonemployers. Since we assume that productivity of entrant and incumbent nonemployers is constant, the average labor input of a nonemployer entrant is the same as the average for nonemployer incumbents, the average size of entrants in any year can be written as:

$$\bar{\ell}_{ent} = \frac{L_{ent}}{N_{ent}} = \frac{L_{ent}^{emp} + \bar{\ell}^{non} \cdot N_{ent}^{non}}{N_{ent}^{emp} + N_{ent}^{non}}, \tag{16}$$

which requires additional information on $\bar{\ell}^{non}$ and N_{ent}^{non}.[7]

To obtain the average labor input of nonemployers $\bar{\ell}^{non}$, we note that in the model, from equations (2) and (3), the share of aggregate revenue generated by a group of firms is equal to its share of aggregate labor, hence, we use the share of aggregate revenue from nonemployers:[8]

$$\bar{\ell}_t^{non} = (agg.\ revenue\ share_t^{non}) \cdot \frac{L_t}{N_t^{non}}. \tag{17}$$

As a side note, we can use this expression and equation (12), to calculate the average productivity of nonemployers relative to all firms and find that it increased from 3.7% in 1982 to 3.8% in 2014, with volatility between 1987 and 2007 similar to that in Figure 4b.

To obtain the number of nonemployer entrants N_{ent}^{non}, we assume a constant exit rate for nonemployers $\lambda^{non} = 0.15$ as reported by Davis et al. (2009) and data on the number of nonemployers over time as follows;[9]

$$N_{ent,t}^{non} = N_t^{non} - (1 - \lambda^{non}) N_{t-1}^{non}. \tag{18}$$

We calculate average firm-level productivity in each year, relative to 1982, using equations (15)-(18), and use it along with the number of firms per worker to calculate the implied aggregate TFP from equation (14). Figure 5a reports aggregate TFP in the extended model compared

[7]Note that the number of employees at employer firms, as reported in the BDS data, does not account for persons engaged that are not formal employees. We adjust average employment by a ratio equal to total persons engaged (from the CPS) over total employees.

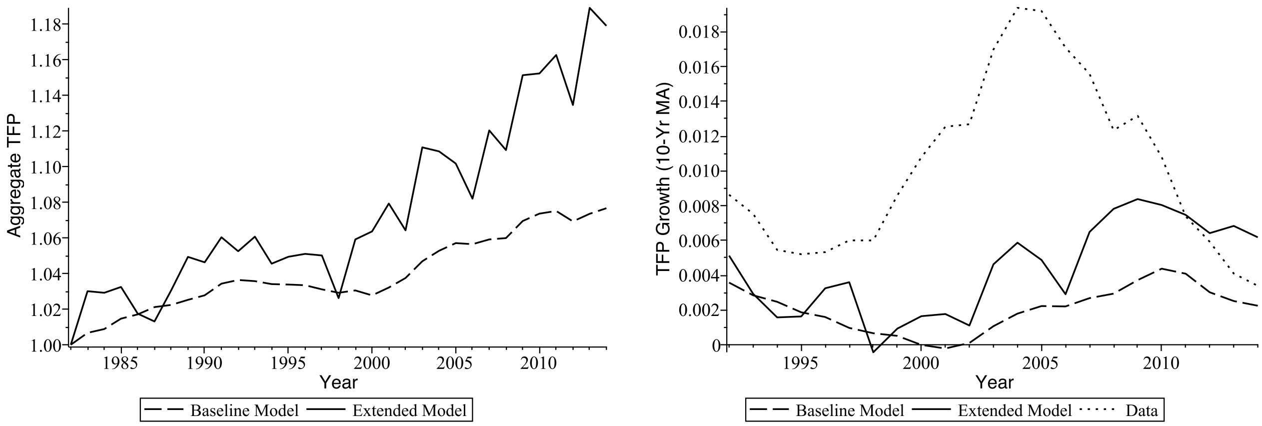
[8]We use Census years starting in 1987. For the years between census years, we assume that the average productivity of nonemployers relative to all firms changes smoothly from one census year to the next. We then calculate the average annual growth rate of this measure between 1987 and 2012, and use it impute average productivity of nonemployers before 1987 and after 2012.

[9]We abstract from nonemployer to employer transitions since we cannot differentiate between new employers and transitioning nonemployers. As a result, we potentially overstate the number of employer entrants, however, we argue this bias should not be quantitatively significant given the small number of transitioning nonemployers in a given year documented by Davis et al. (2009).

Figure 5: Aggregate TFP in Extended Model

(a) Aggregate TFP

(b) Growth in TFP



Notes: TFP growth in the data and implied by the models are reported as 10-year moving averages. TFP data from Fernald (2012), updated to 2014.

with the baseline model. The implied measure of average firm-level productivity (15), equal to the ratio between the two measures of TFP in Figure 5a, is generally increasing from 1982 to 2014. This growth averages 0.11% per year between 1982 and 1997, then a much a higher 0.44% per year from 1997 to 2014. As a result, the cumulative increase in aggregate TFP is higher in the extended model than in the baseline model. From 1982 to 2014, the cumulative increase in TFP is 17.9% compared to 7.7% in the baseline model.

Figure 5b reports the growth rate of aggregate TFP over time in the extended model, reported as a 10-year moving average. From 1982 to 2014, the annual growth averaged 0.52%. The dotted line is the 10-year moving average of observed TFP growth in the U.S. data. The evolution of TFP growth over time implied by the model follows the data well with a correlation 16%, suggesting that business dynamism in fact plays an important quantitative role in driving trends in aggregate productivity. The cumulative growth in TFP observed in the data between 1982 and 2014 amounts to 32%, and as a result the implied measure of TFP in the extended model accounts for 56% of the actual increase. The dashed line in Figure 5b shows a corresponding measure of implied TFP growth in the baseline model with a correlation of -7.5%. Note also that the corresponding implied TFP growth in the baseline model with employer firms not only implies a decline in aggregate TFP over time, but also it is strongly negatively correlated with observed productivity growth with a correlation of -39%.

6 Alternative Drivers of Firm Dynamics

There is an important literature in macroeconomic development linking low average firm size (the ratio of workers to the number of firms) and low aggregate productivity to policy distortions in developing countries. It is therefore relevant to ask whether the increase in the number of firms and hence the decline in average firm size when including nonemployers may be due to changes in distortions rather than changes in exit and productivity growth patterns considered in our model. We consider two broad mechanisms as potential alternative drivers of business dynamism: (a) an increase in resource misallocation and (b) increasing barriers for firms to enter new markets.

Increasing misallocation among employer firms. If larger and more productive firms are effectively taxed at higher rates than smaller unproductive firms, then all firms reduce investment in productivity. This effectively reduces non-production costs for all firms, thereby *increasing* profitability and encouraging entry. In equilibrium, these size-dependent distortions result in more firms that are less productive on average. Hsieh and Klenow (2014) and Bento and Restuccia (2017) show that cross-country differences in the extent to which firm-level distortions are positively related to firm size can go a long way to rationalizing the large differences in average firm size across countries at differing levels of development. We therefore consider whether an increase in the prevalence of idiosyncratic distortions may be driving the increase in the number of firms in the U.S. economy over time.

To assess this mechanism, we use data from the Economic Census for 74 3-digit NAICS industries for the years 2002 and 2012. Within each industry, we collect data on the total number of firms, the number of firms in each size bin, the number of employees per firm within each bin, total payroll, and total revenue.[10] Although we only have data from 2002 to 2012, this period is still characterized by a substantial increase in the number firms of 22%. For each size bin within an industry, we use the ratio of revenue to payroll as our measure of the average distortion faced by each firm within the bin. Hsieh and Klenow (2009) show that, under certain structural assumptions, profit maximization implies that in the absence of distortions, firms within an industry choose labor inputs to equalize the average product of labor across firms within the industry. To the extent that the average product is higher for firms in a large size category relative to a small category, we interpret this as evidence that larger firms face larger implicit taxes.

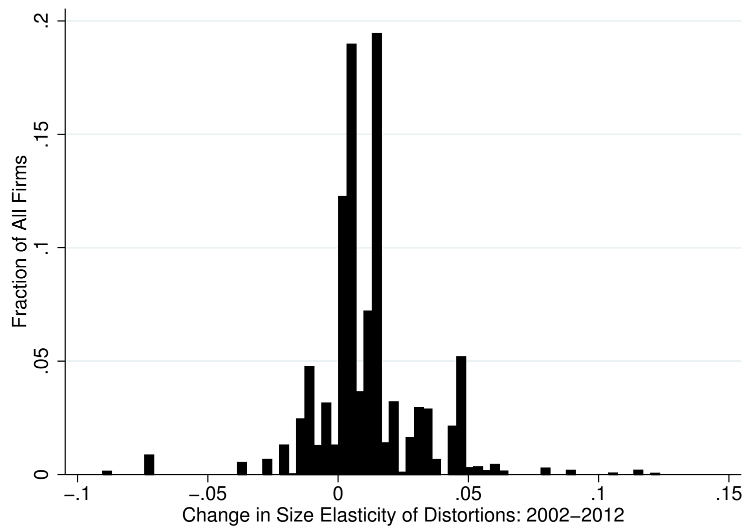
We regress (logged) average products on (logged) employment size within each industry to obtain an estimate of the elasticity of distortions with respect to size as our measure of misallocation.[11] The higher the elasticity, the larger the effective tax rate faced by large firms relative to small firms. For 2002, we find an average elasticity across industries equal to -0.01, with a variance across industries of 0.01. For 2012, we find an average elasticity closer to zero, with a similar variance. These results suggest that there is essentially no change in the extent of misallocation over time as measured by the size elasticity of distortions. Figure 6 reports the fraction of firms in industries that saw a given change in the size elasticity of distortions, from 2002 to 2012. The histogram indicates that the vast majority of firms are in industries that saw little change in this elasticity over time. This finding is consistent with Bils et al. (2020) who use more disaggregated data to account for measurement error and show that there is no significant change in misallocation over time in the U.S. economy. We conclude that increased misallocation is not driving the increase in the number of firms in the U.S. economy over time.

Increasing misallocation between employer and nonemployers. The previous evidence relies on measures of distortions inferred from employer data, as a result it abstracts from potential distortions affecting firm's decisions whether to have employees or not. We now

[10] For manufacturing industries, we use establishments rather than firms, and value added rather than revenue.

[11] We exclude all firms with less than 5 employees to address the potential issue of unpaid workers not being counted in small firms, hence number of firms is defined as the number of firms with at least 5 employees.

Figure 6: Distribution of Changes in the Size Elasticity of Distortions



turn to assess this possibility. Consider that all potential entrants must pay an entry cost to become a producer, but now firms that hire labor (in addition to labor from owner-managers and/or informal workers) must incur an additional fixed cost c_{emp}.[12] Further assume that nonemployers use (informal) labor less productively than employer firms use (formal) labor, so that nonemployers have an effective productivity of $(\psi z)^{1-\alpha}$, $\psi < 1$. The additional fixed cost of becoming an employer firm encourages firms which would otherwise hire a small number of employees to remain nonemployers if the additional profit from these workers is not high enough to cover the effective marginal cost of labor. If c_{emp} increases over time, other things equal, then the fraction of firms choosing to become nonemployers increases. This mechanism could therefore be a potential driver of the increase in the fraction of firms that are nonemployers over time. Note that an increase in c_{emp} over time would also decrease aggregate productivity, as a higher fraction of firms would be subject to the cut in productivity ψ. This suggests a potential mechanism for aggregate productivity gains from an increase in the number of firms to be lower than in the baseline model. This mechanism, however, generates two counter-factual implications. First, we should observe a gap in the firm-level employment distribution, with no employer firms reporting (for example) one employee. And this gap should be increasing over

[12]Note the fixed cost could be a tax on wages that is affected by regulation and policy.

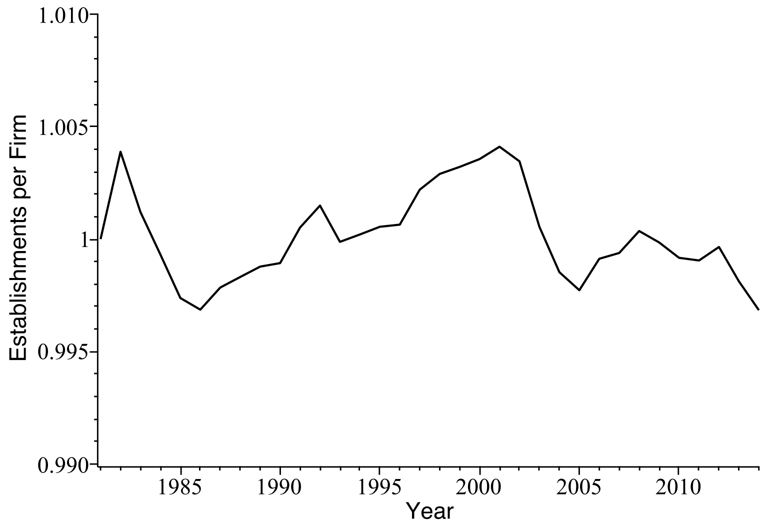
time. There has been no evidence of such a gap at any point in time, nor of a growing gap over time. Second, a higher c_{emp} would generate a larger divergence over time between the change in the fraction of firms that are nonemployers and the aggregate revenue share of nonemployers. To see this, we express the revenue share of nonemployers in the following way;

$$\text{RevShare}_{non} = \left(\frac{N_{non}}{N_{all}}\right)\frac{\psi \mathbb{E}_{non}(z)}{\bar{z}_{\psi}},$$

where $\mathbb{E}_{non}(z)$ denotes average z across all nonemployers, and $\bar{z}_{\psi}$ denotes average z across all firms adjusted by ψ for nonemployers. If the underlying distribution of productivity across firms is constant over time, then our baseline model (where $\psi = 1$) predicts that the nonemployer revenue share should grow faster than the fraction of firms that are nonemployers. This is because an increase in nonemployers implies an increase in the threshold productivity above which firms are counted as employers, thereby increasing the average productivity of nonemployers. As a result, the baseline model over-predicts the increase in the revenue share of nonemployers from 1987 to 2012 while under-predicting the increase in the share of firms that are nonemployers. If nonemployers face a lower effective productivity (due to $\psi < 1$), then an increase in the fraction of firms that are nonemployers should still result in the same increase in the average effective productivity of nonemployers as implied by the baseline model. But because a higher fraction of firms are effectively less productive, average effective productivity across all firms ($\bar{z}_{\psi}$) drops, unlike in the baseline model. This implies that the revenue share of nonemployers must increase even more than in the baseline model, for any given increase in the fraction of firms that are nonemployers. This discussion suggests that an increase over time in the cost of becoming an employer firm is not driving observed trends.

Increasing entry barriers to new markets. Another possible alternative explanation for the increase in the number of firms is higher barriers to entering a market. Bento (2020) shows that when firms choose how many markets to enter, and the cost of entering is increasing in the number of markets entered, barriers to market entry (distinct from barriers to starting a firm)

Figure 7: The Number of Establishments per Firm (Relative to 1981)



encourage more firm startups, with each firm competing in fewer markets in equilibrium. As a result, each market is characterized by fewer competing firms, and aggregate productivity drops even as the aggregate number of firms increases. If barriers to market entry have been increasing in the U.S. economy, we should observe fewer firms competing in each local market, even as the aggregate number of firms increases. We do not have data on the number of firms present in each market, and defining a market is difficult. But we can consider how the aggregate number of establishments changes over time. To the extent that firms create multiple establishments to access multiple markets, the number of establishments per firm can serve as a proxy for the number of markets per firm. Figure 7 reports that the number of establishments per firm essentially remained constant from 1981 to 2014, even as the number of firms per worker grew by 53%. Note that Aghion et al. (2020) and Cao et al. (2020) make a similar observation about establishments per firm over time in the context of employer firms. This suggests that increasing barriers to market entry are not likely driving the increase in the number of firms per worker in the U.S. economy.

In summary, the evidence presented in this section suggests that various mechanisms encouraging firm entry while lowering TFP, which are prevalent in developing countries, are not driving the increase in business dynamism in the United States, that is the increase in the total number of firms and the share of nonemployer firms.

7 Conclusions

An important literature documenting a decline in business dynamism in the U.S. over the last several decades has focused solely on employer firms. We consider a broader measure of firms that includes nonemployers, and find that the total number of firms has diverged dramatically from the number of employer firms over time. We interpret this fact, along with the evolution of the employment distribution across firms, through the lens of a model of firm dynamics based on Hopenhayn (1992). We show that accounting for nonemployers drastically changes the implications for aggregate productivity. Although nonemployers are small relative to employers, the increase in the number of firms and in firm-level productivity together imply that business dynamism has been responsible for over one half of observed aggregate productivity growth from 1982 to 2014. This is in striking contrast to a decrease in productivity implied by a model considering only employer firms.

Decker et al. (2016) and Li (2017) show that standard measures of business dynamism, which focus on employer firms, do not correlate well with TFP growth, casting doubt on the quantitative importance of theories of firm dynamics. We show that our broader measure of business dynamism, which accounts for nonemployers and the evolution of the size distribution over time, follows TFP growth in the data more closely since the 1980s. Nevertheless, the question remains what then accounts for the productivity slowdown in the U.S. economy in recent decades. Our results indicate it is not a decline in business dynamism. One promising recent study instead focuses on the decline of the quality of innovative activity resulting from a misallocation of R&D investments across firms (Ayerst, 2020), driven by firm-level heterogeneity in the wedge between the private and social return to innovation.

Our results suggest several avenues for future research. It would be useful to relate our comprehensive measure of the number of firms with recently documented trends in market concentration and price-cost markups, as documented in De Loecker et al. (2020), and Rossi-Hansberg et al. (forthcoming). Relatedly, theories developed to explain increasing markups and market concentration, as well as the declining labor share of aggregate income, have taken as given a decline in the number of firms. For instance, Akcigit and Ates (2019) relate these trends to declining business dynamism. As a result, an important direction for future research may be exploring mechanisms that can account for these trends in the context of *higher* firm entry. We have abstracted from the underlying causes of changes in exit rates and productivity growth across firms and over time. Understanding these patterns remains an important area of research (Aghion et al., 2020; Cao et al., 2020). Similarly, given the growing importance of nonemployer firms in the U.S. data, it is essential to document and better understand the nature of nonemployer business activity. We leave these important explorations for future research.

References

Acs, Zoltan J., Brian Headd, and Hezekiah Agwara, "Nonemployer Start-up Puzzle," Technical Report, Small Business Administration, Office of Advocacy 2009.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," 2020. Unpublished manuscript.

Akcigit, Ufuk and Sina T. Ates, "What Happened to U.S. Business Dynamism?," *NBER Working Paper No. 25756*, 2019.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," *Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Ayerst, Stephen, "Innovator Heterogeneity, R&D Misallocation, and the Productivity Growth Slowdown," 2020. Unpublished manuscript, International Monetary Fund.

Bento, Pedro, "Competition, Innovation, and the Number of Firms," *Review of Economic Dynamics*, 2020, *37*, 275–298.

_ **and Diego Restuccia**, "Misallocation, Establishment Size, and Productivity," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 267–303.

— **and** — , "On Average Establishment Size across Sectors and Countries," *Journal of Monetary Economics*, forthcoming.

Bils, Mark, Peter J. Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," *NBER Working Paper No. 26711*, 2020.

Bollard, Albert, Peter J. Klenow, and Huiyu Li, "Entry Costs Rise with Development," 2016. Unpublished manuscript, Stanford University.

Cao, Dan, Henry R. Hyatt, Toshihiko Mukoyama, and Erick Sager, "Firm Growth Through New Establishments," 2020. Unpublished manuscript.

Davis, Steven J., John Haltiwanger, Ronald S. Jarmin, C. J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky, "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes," *Producer Dynamics: New Evidence from Micro Data*, 2009, pp. 329–366.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda, "Declining Business Dynamism: Implications for Productivity," *Hutchins Center Working Paper No. 23*, 2016.

Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda, "The Secular Decline of Business Dynamism in the United States," 2014. Unpublished manuscript, University of Maryland.

Fairlie, Robert W., Javier Miranda, and Nikolas Zolas, "Job Creation and Survival among Entrepreneurs: Evidence from the Universe of U.S. Startups," 2018. Unpublished manuscript, University of California, Santa Cruz.

Fernald, John G., "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *FRBSF Working Paper No. 2012-9*, 2012.

Furman, Jason and Peter Orszag, "Slower Productivity and Higher Inequality: Are They Related?," *PIIE Working Paper No. 18-4*, 2018.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda, "Who Creates Jobs? Small versus Large versus Young," *Review of Economics and Statistics*, 2013, *95*, 347–361.

Hopenhayn, Hugo A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60*, 1127–50.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania, "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," *NBER Working Paper No. 25382*, 2019.

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

_ **and** _ , "The Life Cycle of Plants in India and Mexico," *Quarterly Journal of Economics*, August 2014, *129* (3), 1035–1084.

Hyatt, Henry, Seth Murray, and Kristin Sandusky, "Business Income Dynamics and Labor Market Fluidity," 2020. Unpublished manuscript.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin, "Demographic Origins of the Startup Deficit," 2019. Unpublished manuscript.

Li, Huiyu, "How Does Business Dynamism Link to Productivity Growth?," *FRBSF Economic Letter No. 2017-01*, 2017.

Pugsley, Benjamin and Ayşegül Şahin, "Grown-Up Business Cycles," *Review of Financial Studies*, 2019, *32*, 1102–1047.

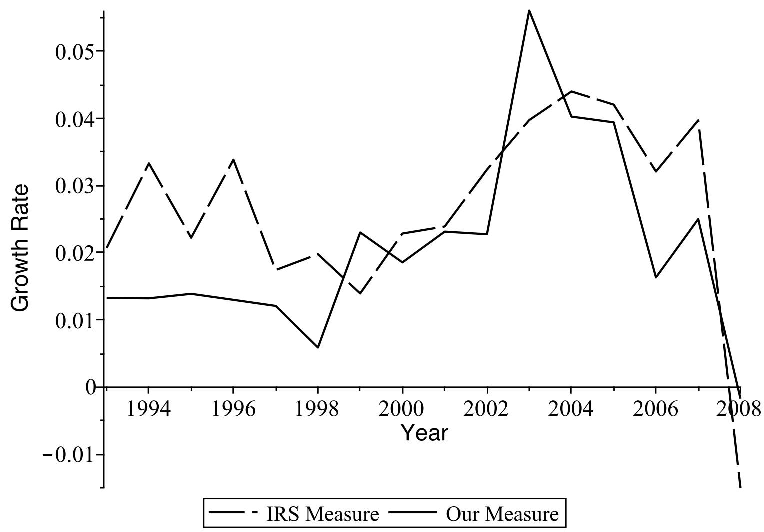
Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration," *NBER Macroeconomics Annual*, forthcoming.

Appendix

A Data Imputation

For the years between 1981 to 1991, we impute the number of nonemployers by using the growth rate in the total number of firms reported by the IRS (constructed using tax returns). We work backwards from 1992, imputing the total number of firms using the growth rate in each year from the IRS data, then subtracting the number of employers (from BDS) to obtain the number of nonemployers. Figure 8 documents the fact that the growth rate in the total number of firms reported by the IRS tracks very closely our measure of the growth rate of the total number of firms over the years for which we have data for nonemployers and employers. Hence, we argue this imputation of the number of nonemployer firms is reasonable.

Figure 8: The Growth Rate in the Number of Firms, IRS and Our Measure



B Sectoral Composition in Number of Firms

Given the structural transformation in the U.S. economy over the last several decades, it is important to assess whether the large increase in the total number of firms per worker is driven by within-sector changes in net entry or by changes in sectoral employment shares, that is changes from sectors with a low number of firms per worker to sectors with a high number of firms per worker. We analyze how sectoral employment shares have evolved over time between 1983 to 2014 for 9 sectors of the economy: agriculture, forestry, and fishing; mining;

construction; manufacturing; wholesale trade; retail trade; transportation, communication, and utilities; finance, insurance, and real estate; and other services. We find that the most significant change is the reallocation of employment away from manufacturing to other services. Within these two sectors, firms per worker in manufacturing rose by 47%, while firms per worker in other services rose by a close 46%, which suggests that the process of structural transformation is not driving the increase in the total number of firms per worker. Indeed, firms per worker rose in seven out of nine sectors. The only sectors that experienced a drop in the number of firms per worker are Mining (−52%) and Retail Trade (−16%).

Table 1: The Role of Structural Transformation in Total Firms per Worker

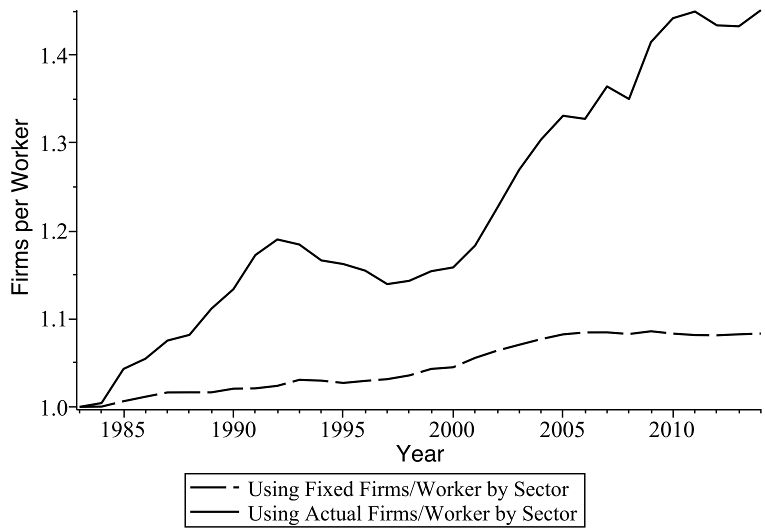
	Employment share (%)		Firms per worker (×100)	
	1983	2014	1983	2014
Sectors				
Agriculture, forestry, and fishing	4	2	9	16
Mining	1	1	25	12
Construction	7	7	21	29
Manufacturing	20	11	3	4
Wholesale trade	4	3	17	20
Retail trade	12	12	21	17
Transportation, communication, and utilities	6	6	11	24
Finance, insurance, and real state	7	7	29	38
Other services	40	52	15	22
Aggregate	100	100	14	20

Nevertheless, to get a more concrete quantitative assessment of the importance of structural transformation to the increase in the number of firms per worker, we compute a counterfactual aggregate number of firms per worker assuming that the number of firms per worker in each sector is fixed at 1983 levels. Changes in this counterfactual measure over time are therefore solely driven by changes in sectoral employment shares. Figure 9 reports this counterfactual measure of the aggregate number of firms per worker, along with the actual number of all firms per worker for comparison. The counterfactual shows that only 20% of the increase in the total number of firms per worker can be accounted for by structural change.

C Alternative Beliefs in the Baseline Model

We show that our assumptions about the beliefs of firms in the baseline model affect the implied evolution of the entry cost over time, but do not affect our calculation of the implied evolution

Figure 9: Total Number of Firms per Worker, Actual vs. Counterfactual



Notes: The solid line represents the evolution of the total number of firms per worker in the data, whereas the dashed line is the counterfactual evolution of the total number of firms per worker when firms per worker in each sector is kept fixed at 1983 levels.

of aggregate productivity. In the baseline model we assume firms always believe the current supply of labor L and entry cost c_e will persist indefinitely, and are subsequently shocked each period. We now assume that firms know the future paths of both L and c_e with certainty. For this exercise we assume that the economy is in a steady state in 1981, such that L and c_e were previously constant at 1981 levels. We further assume these variables stop changing after 2014.

Per-period optimal output, labor demand, and profits are still described by equations (2) through (4), as functions of firm productivity. The wage is still described by equation (5), as a function of the number of firms per worker. That equation (5) still holds in each period implies that our calculation of implied TFP (11) also holds, given the observed number of firms per worker in each year in the data.

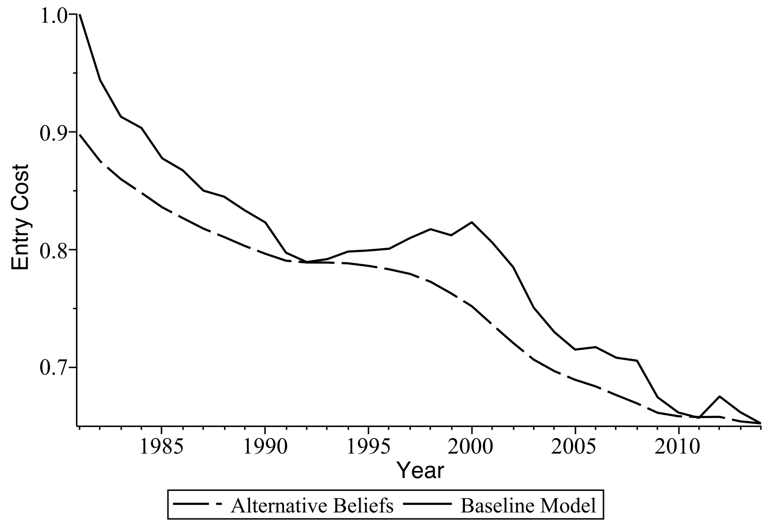
Free entry requires that the expected discounted profits of entrants are exactly equal to the entry cost in the period they enter. This can be expressed recursively as;

$$c_{e,t} = \frac{(1-\alpha)}{(N_t/L_t)} + \rho c_{e,t+1}, \tag{19}$$

$$c_{e,2014} = \frac{(1-\alpha)}{(N_{2014}/L_{2014})(1-\rho)}. \tag{20}$$

Potential entrants now take into account future entry and labor supply growth when making their entry decision. Taking as given the number of firms per worker in the data, the more firms per worker in the future, the lower the current entry cost must be to rationalize a given

Figure 10: Implied Entry Cost with Alternative Beliefs



current number of firms per worker. In Figure 10 we show how the implied entry cost c_e must evolve over time in order to match the observed evolution of the number of firms per worker in the data. Compared to the baseline, this alternative implied entry cost must be lower in 1981, given that firms are now taking into account the future observed increase in the number of firms per worker. By 2014, the two measures converge, as must be the case since the two sets of beliefs also converge. With these alternative beliefs, the implication for the evolution of entry costs is much the same, except here the implied entry cost is less volatile over time.

D Share of Nonemployers in Baseline Model

To assess the ability of our baseline model to explain the change in the number of nonemployers over time, given observed changes in the total number of firms, we calibrate the distribution of productivity across firms $G(z)$ to match U.S. data for 1987. The BDS data includes the number of employer firms falling within 12 different employment-size bins, along with average employment per employer firm within each bin. Two issues arise here. First, we do not have data for labor used in nonemployer firms (ie., owners, informal workers). Second, the BDS data does not report informal workers and owners in employer firms. To address these issues we use data from the Survey of Business Owners and the Statistical Abstract of the U.S., from which we can calculate the share of aggregate revenue generated by nonemployer firms. This data is available in census years starting in 1987.

We start by noting that in the model the share of aggregate revenue generated by a group of firms is equal to its share of aggregate labor, from (2) and (3). We therefore assume that the labor used across all nonemployers is equal to the share of revenue generated by nonemployers. For our purposes here we need a distribution of labor across nonemployers and across small employers (those with less than 5 formal employees). We make a further assumption about the employment data – that the difference between total persons engaged (from the CPS) and formal employees (from the BDS) is distributed across all firms in proportion to their reported number of employees. In other words, if firm A is reported to have 5 times the employees of firm B, then we assume firm A also has 5 times the number of total workers as firm B. In 1987 nonemployers are therefore assumed to employ 2.8% of aggregate labor, while small employers employ 6.2%. Taking into account the numbers of each type of firm, these translate to an average employment size of 0.24 for nonemployers and 2.85 for small employers.

To map the above results to relative firm-level productivity, we note that equation (2) implies the following productivity ratio between two arbitrary firms i and j with different levels of employment;

$$\frac{z_i}{z_j} = \frac{\ell_i}{\ell_j}. \tag{21}$$

To infer the distribution of productivity across nonemployers and small employers, we assume nonemployers are firms with optimal (reported) labor less than 1, which translates to optimal actual labor less than 1.24. For comparison, small employers are firms with actual labor between 1.24 and 6.2 (1-5 multiplied by 1.24). We assume the distribution of productivity within each group is described by the Beta distribution, Beta$(a, 1|z_L, z_U)$, where z_L and z_U are the bounds of each distribution, $z_L = 0$ for nonemployers, and a is specific to the group.[13] The flexibility of this distribution allows us to match the average z (average labor) for each group from the data. Doing this, we obtain $a_{non} = 0.24$ and $a_{small} = 0.48$.[14] Given $G(z)$, we then calculate the fraction of firms that are nonemployers generated by the model as the number of firms changes each year (again, we assume the entry cost changes over time to generate this outcome).

[13] The pdf of this distribution is $f(z \mid z_L < z < z_U) = \frac{a(z-z_L)^{a-1}}{(z_U-z_L)^a}$.

[14] We only need within-group distributions of productivity for nonemployers and small employers, as the share of these firms in 1987 is 88%, while the share of nonemployers never rises above 82.5%.