NBER WORKING PAPER SERIES

THE ROLE OF NONEMPLOYERS IN BUSINESS DYNAMISM AND AGGREGATE PRODUCTIVITY

Pedro Bento
Diego Restuccia

Working Paper 25998
http://www.nber.org/papers/w25998

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2019, Revised July 2024

For helpful comments we thank the editor, two anonymous referees, Benjamin Bridgman, John Haltiwanger, Pete Klenow, Benjamin Pugsley, Pavel Sevcik, and audiences at the 2023 Canadian Macroeconomics Study Group Meeting, 2021 Society for Economic Dynamics Meeting inf Minnesota, Texas A&M University, University of Texas at Austin, Utah State University, 2020 Econometric Society North American Winter Meeting in San Diego, 2019 Midwest Macroeconomic Meeting at Michigan State, and 2019 RCEA Growth, Innovation, and Entrepreneurship Conference. 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. The views expressed herein are not necessarily those of the Bank of Canada or the National Bureau of Economic Research and are the authors' alone.

The Role of Nonemployers in Business Dynamism and Aggregate Productivity
Pedro Bento and Diego Restuccia
NBER Working Paper No. 25998
June 2019, Revised July 2024
JEL No. E02,E1,O1,O4,O51

ABSTRACT

The well-documented decline in business dynamism, measured by the net entry rate of employer firms, has been proposed as an explanation for the productivity growth slowdown in the United States. We examine the role of nonemployers, firms without paid employees, in business dynamism and aggregate productivity. Between 1982 and 2014, the total number of firms per worker including nonemployers increased by 41%, whereas employer firms per worker declined by 8.7%. Using a standard model of firm dynamics, we derive the implications for aggregate productivity associated with changes in firms per worker and relative size distributions. We find that firm dynamics imply an increase in aggregate productivity of 15.6%, about half the growth observed in the data, that is equally shared by the changes in firms per worker and relative sizes. These results contrast markedly with the much weaker 2.1%growth in aggregate productivity from firm dynamics when abstracting from nonemployer firms. Our results suggest the productivity growth slowdown is not due to changes in net firm entry, and highlight the quantitative importance of comprehensive measures of business dynamism in the U.S. data.

Pedro Bento
Texas A&M University
1101 Hollow Creek Dr. Unit 2208
Austin, TX 78704
pbento@tamu.edu

Diego Restuccia
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
and NBER
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 growth in aggregate total factor productivity (TFP) before the Great Recession. In this paper, we construct a broader measure of the total number of firms that contributes to a more comprehensive characterization of business dynamism in the United States, and 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 widely 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 show using a standard model of firm dynamics to draw TFP implications from firm counts or startup rates, that inferences are substantially biased if only a subset of all firms is counted.

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 1982 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 −8.7% from 1982 to 2014, consistent with the findings in Karahan et al. (2019), the total number of firms per worker in our measure with nonemployers increased by 41% over the same period.

We consider a standard model of firm dynamics building from Hopenhayn (1992) in order to quantify the impact of the surge in the number of firms on aggregate productivity. A single homogeneous good is produced by firms every period. Firms have access to a decreasing returns to scale technology and are heterogeneous with respect to their total factor productivity. Firms operate in competitive output and input markets. A large number of potential entrants can become producers by paying a resource cost and draw their productivity from a constant distribution function after entry. Every period, incumbent firms face exogenous productivity growth and exit that depend on their age and productivity and can vary over time. The economy features exogenous time-varying processes for economy-wide productivity, labor supply, the real interest rate, and the relative entry cost. Firms know the evolution of these exogenous variables in making entry decisions. We further assume that potential entrants believe that exit rates and productivity growth rates are equal to their cross-section averages and know their evolution over time.

We characterize the competitive equilibrium solution of the model and show that with suitable processes for the exogenous variables, the model matches the time series data for the total number of firms per worker, the increasing share of nonemployers, and the reallocation of economic activity across firms. The model implies a substantial decline in entry costs relative to output per worker, which is consistent with some evidence (Koske et al., 2015; Bollard et al., 2016); and the reallocation of economic activity is also consistent with recent studies of market concentration (Autor et al., 2020; Rossi-Hansberg et al., 2021). We then use the model to examine the implied growth in aggregate productivity associated with firm dynamics and contrast the results with those abstracting from nonemployers.

The calibrated model implies that measured aggregate TFP associated with firm dynamics increased by 15.6% from 1982 to 2014. As a reference for comparison, measured aggregate TFP in the data during the same time period increased by 32% suggesting that TFP growth associated with firm dynamics contributed substantially to the productivity growth in the U.S. economy. Decomposing the growth in measured aggregate TFP from firm dynamics in the model, we find that about half of the increase (7.1%) is due to the change in the number of firms per worker, with the remaining half (8%) due to the reallocation in economic activity (the shift in the relative firm size distribution).

We contrast these results with those implied by firm dynamics when abstracting from nonemployers. The implied cumulative growth in measured aggregate TFP from firm dynamics between 1982 to 2014 is only 2.1%, compared to 15.6% when accounting for nonemployers. The lower aggregate TFP growth associated with employers arises mostly due to a decline in employers per worker, generating a cumulative -4.8% drop in aggregate TFP (versus 7.1% increase with nonemployers), which is offset by an increase in aggregate TFP of 6.9% due to rising average productivity $\bar{z}$ (versus 8% increase with nonemployers). Accounting for nonemployers implies a much larger growth in measured TFP arising mostly due to the increase in firms per worker, but also due to an increase in resource reallocation driving higher average firm-level productivity. These contrasting results between firm dynamics characterized with and without nonemployers highlight the quantitative importance of a more comprehensive measure of business dynamism in the U.S. data. We also use the model to examine the factors leading to the increase in the share of nonemployers over time, finding that both the increase in the total number of firms per worker and the increase in average firm-level productivity are important for this outcome.

Our more 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 persistent decline in net entry rates for employer firms. We show that the net entry rate of all firms declined until the mid-1990s, then began increasing. Hopenhayn et al. (2019) and Karahan 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, net firm entry has contributed substantially to productivity growth. Hopenhayn et al. (2019) emphasizes 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 net firm entry 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 on business dynamism 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. Our analysis does not account for changes over time in job reallocation rates or labor flows that might affect the effective productivity of the workforce. 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. Nevertheless, to the extent that lower rates of job destruction are driven by lower exit and higher growth rates for the most productive firms, our model suggests lower destruction rates may be accompanying the higher aggregate productivity driven by the shift in employment towards these firms.

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, thereby lowering aggregate TFP (Guner et al., 2008; Hsieh and Klenow, 2014; Bento and Restuccia, 2017, 2021; Bento, 2020). We assess whether changes in misallocation may be driving the trend in the number of firms in the United States. Using limited data on employment and revenue across employment-size categories, we do not find substantial evidence of increased misallocation among employer firms.

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 a standard model of firm dynamics augmented to include nonemployers in order to assess the quantitative impact of firm dynamics on aggregate productivity. Section 5 provides a discussion of potential alternative drivers of business dynamism. We conclude in section 6.

2 Nonemployer U.S. Businesses

We construct a comprehensive measure of the total number of (non-farm) 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. Nonemployers are businesses with no paid employees, including self-employed entrepreneurs, that have annual business receipts of $1,000 or more ($1 or more in construction) and are subject to federal income taxes. In 2012, 86% of nonemployers were sole proprietors such as real estate agents, independent contractors, and small businesses with informal/family workers, 7% were partnerships, and 7% were corporations. A comprehensive measure of firms may be relevant in understanding changes in net entry rates over time, as it is the case when considering very small firms in the context of cross-country differences in establishment size (Bento and Restuccia, 2017, 2021).

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 firms is similar to that of small employers. 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 as 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 do not 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 9% of all employers exit each year between 1994-1997, the corresponding exit rate for all nonemployers is higher, about 15% (Davis et al., 2009). The smallest employers (less that 5 employees) exit at a very similar rate of 13.5% between 1994-1997.[1]

We note that most owners of nonemployer firms also work in other jobs while running their nonemloyer firms. For example, in 2017 only about one third of all entrepreneurs identified self-employment as their main occupation according to Current Population Survey data, indicating fewer hours spent working at their own business than working for another business. See also Abraham et al. (2019) for a discussion of the distinction between measures of self-employment inferred from the Current Population Survey and Nonemployer Statistics.

From these facts, and absent more detailed data on nonemployers, we conclude that a reasonable starting approach to modeling nonemployer businesses is to assume that they are similar to small employers with higher exit rates and operating at a lower scale, possibly because of lower productivity. This interpretation follows how small employers are treated relative to larger employers in the literature. We distinguish nonemployers from employers as businesses demanding less than one unit of labor, with labor supplied by owner-maganers who may also supply labor to other firms, while labor of employers is supplied by owners and hired labor. In Section 4, we consider this approach in extending a standard model of firm dynamics to include nonemployers, and discuss the implications of alternative assumptions about the distinction between nonemployers and employers in Section 4.4.

[1]Employer exit rates are taken from the U.S. Census Bureau's Business Dynamics Statistics data.

3 Data

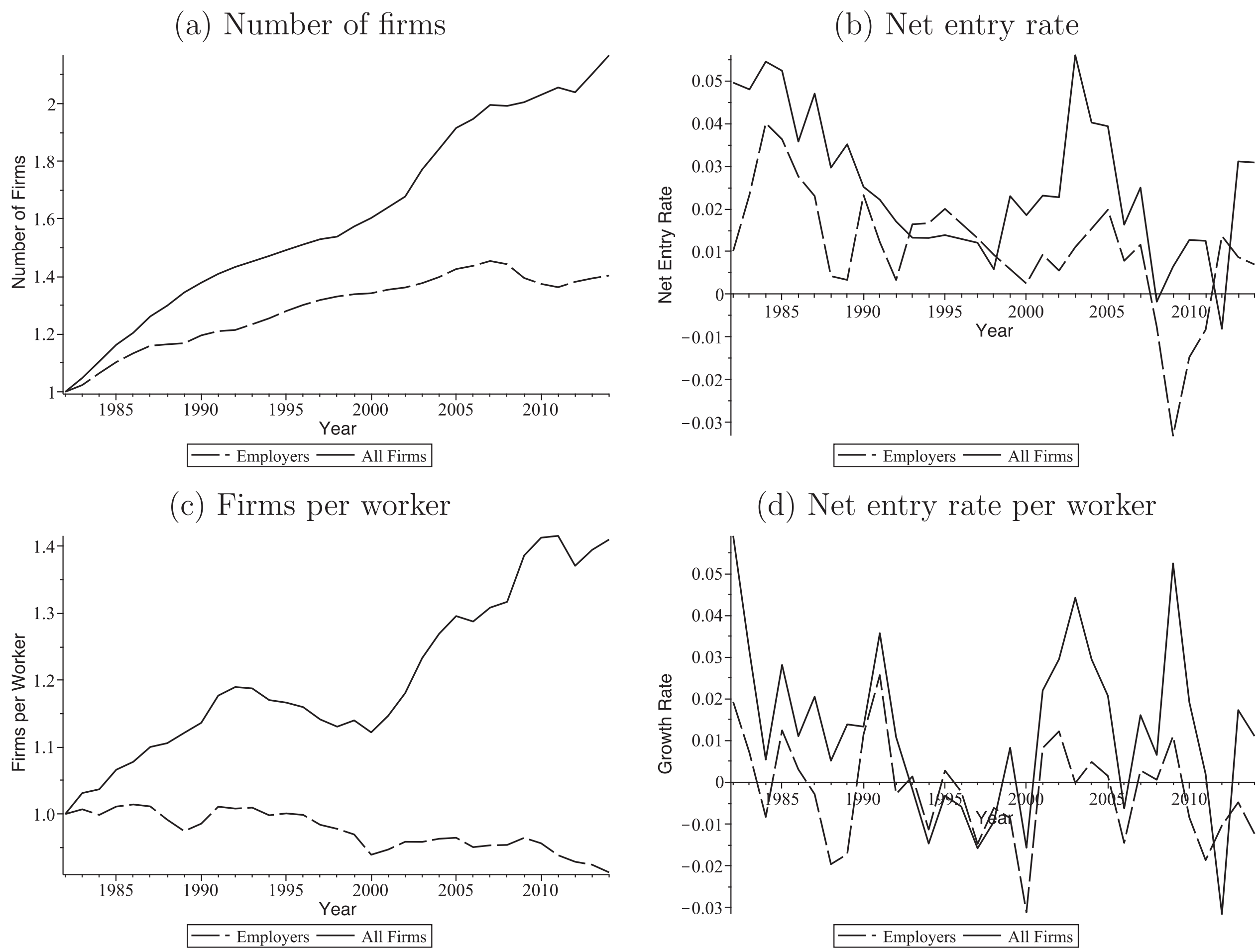
We describe the data and procedure used to construct our measure of the total number of firms over time in the U.S. economy. Data for employer firms is from the U.S. Census Bureau's Business Dynamics Statistics (BDS), the standard data source in the business dynamism literature (Decker et al., 2014). The data for employers contain 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 are 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 1982. 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 1982 to 2014, while the number of employers increased by 40%, the total number of firms increased by a striking 117%. Note that although we have data for 1981, our analysis requires changes in variables, as a result, for consistency we focus on the period from 1982 to 2014.

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 1982. 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. Data from Nonemployer Statistics (NES), Business Dynamics Statistics (BDS), and Statistical Abstract, U.S. Census Bureau, and Current Population Survey (CPS), U.S. Bureau of Labor Statistics.

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 (minus government workers) 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. The CPS data includes farm employment, while our firm data excludes farms. We address this inconsistency in Appendix B, where we show our results are unaffected if we adjust aggregate employment to remove farm workers. 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 1982 to 2014, the number of firms per worker increases by 41%, whereas the number of employers per worker decreases by 8.7%. Our measure of CPS employment is consistent with the measures of labor force participation analyzed in Karahan et al. (2019) and Hopenhayn et al. (2019).

A concern with changes in nonemployer counts over time relates to the criteria used to include a business into the NES data. In Appendix C we discuss how several criteria might bias our firm counts. We focus on a minimum revenue threshold that businesses without paid employees must surpass to be included. For most sectors, this threshold has remained fixed over time in nominal terms at $1,000. This implies that some of the increase in the measured number of nonemployers just discussed might be the result of a decreasing inflation-adjusted revenue threshold over time. To examine this potential issue, we use available data for the Census 2012 year, where we have the proportion of nonemployer businesses that earned between $1,000 and $5,000 in revenue, the lowest and narrower category of nonemployer businesses reported. Using the GDP deflator between 1982 and 2012, we calculate that $1,000 in 1982 represent $2,036 in 2012. This implies that about one quarter of the range in the lowest revenue category may be affected. As a result, assuming a uniform distribution of nonemployer businesses in the lowest category, we estimate that one quarter of the businesses in the lowest revenue category were existing businesses in 1982 that were not counted and removing these businesses implies that the growth in the total number of firm between 1982 and 2012 is 31 percent compared with the 37 percent in our data. That is, this adjustment implies that the growth in the total number of firms per worker could be 5 percentage points lower in 2012 or only about 15% of the total growth (see Appendix C for more details). We conclude that while the nominal revenue inclusion issue does affect the counts of nonemployer businesses, quantitatively it does not appear to undermine the substantial reported growth in nonemployer businesses during our sample period.

An alternative calculation to assess the accuracy of our baseline nonemployer counts with respect to the nominal revenue threshold is to examine counts of sole proprietors over time in the IRS tax-return data, which do not impose any inclusion threshold. Recall that most nonemployer businesses are sole proprietors. The data are available from 1997 to 2014 but note that it includes both nonemployer and employer sole proprietors. We compare the growth rates in sole proprietors per worker during the 1997-2014 period in the IRS data and in our data. We find almost identical growth rates: 25.0% in the IRS data and 24.6% in our data.[2] This evidence further reinforces our previous conclusion that it does not appear that the fixed nominal revenue threshold is driving most of the growth in nonemployer businesses in our data.

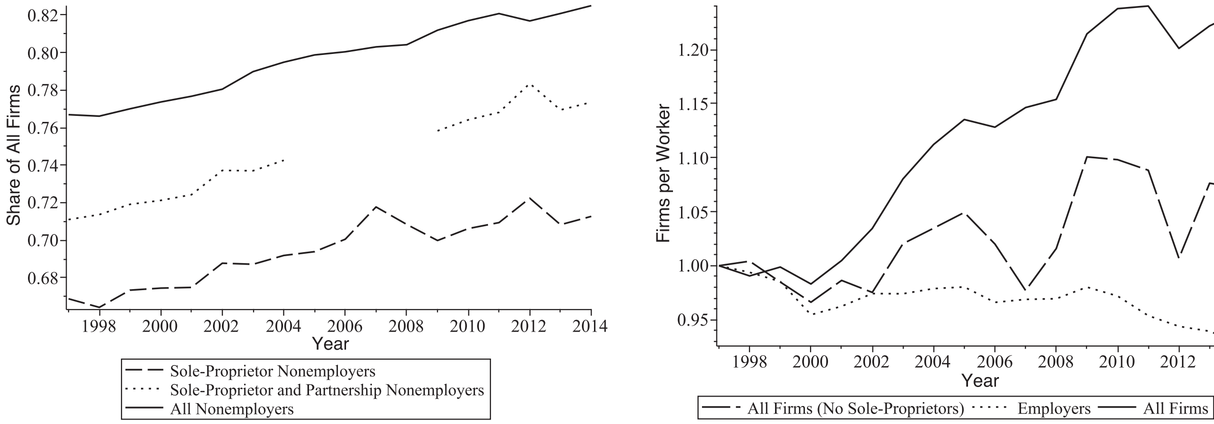
Another potential concern with nonemployer counts is that it may include individuals that register a business for taxation purposes, while effectively working for an employer firm. These types of nonemployers would be categorized in the data as sole-proprietors. The striking difference in the number 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 2a illustrates how the composition of nonemployers with respect to their legal form evolved from 1997 to 2014. As a fraction of total nonemployers, the share of sole-proprietors is 87.2% in 1997 and 86.4% in 2014. As a result, the number of partnerships and incorporated nonemployers increased even more than did all nonemployers. As a share of all nonemployers, partnerships and incorporated nonemployers grew from 12.8% in 1997 to 13.6% in 2014. Figure 2b compares the number of all firms per worker, the number of firms

[2]We obtain the number of employer sole proprietors by multiplying the total number of employers in 1997 and 2014 by the fraction of employers that are sole proprietors in each year. These fractions are from the U.S. Census' Survey of Business Owners data for 1997, and the U.S. Census' Statistics of U.S. Businesses for 2014.

Figure 2: Composition of Nonemployers and All Firms

(a) Composition of Nonemployers

(b) Firms per Worker with and without Sole-Proprietor Nonemployers



Notes: Panel (a): The solid line represents the total number of nonemployers as a share of all firms in the data, the dotted line represents the share of all firms that are sole-proprietor and partnership nonemployers, and the dashed line represents the share of all firms that are sole-proprietor nonemployers. Panel (b): The solid line represents the total number of firms per worker in the data, the dotted line is employer firms per worker, and the dashed line is the total number of firms per worker excluding nonemployer sole-proprietors. In each case, levels are normalized to 1 in 1997. Data for both panels from Nonemployer Statistics (NES) and Business Dynamics Statistics (BDS), U.S. Census Bureau, and Current Population Survey (CPS), U.S. Bureau of Labor Statistics.

without counting 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-proprietor nonemployers, a 7% increase rather than a 24% increase with all firms, it is still markedly higher than the −7% for employers per worker.

We also note that 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 D 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 one fifth of the increase in the number of firms per 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 somewhat over the last three decades, the number of all firms per worker that includes nonemployers has risen substantially by 41%. In the next section, we develop a framework to examine the implications of this growth in the total number of firms for aggregate productivity and other variables of interest.

4 Model

We consider a version of the firm dynamics model in Hopenhayn (1992) in order to provide a mapping from changes in the net entry of firms and in the distribution 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. We aim to make the model as rich and flexible as possible, allowing for multiple factors to change over time. Nevertheless, our analysis is restricted by the limited data available for nonemployers. Given this limitation, our approach is parsimonious with respect to the modeling of nonemployers, in particular, we assume nonemployers have access to the same technology than employers but whose demand for labor is less than one unit, which is supplied by the business owner. This approach is consistent with the relative sizes of nonemployer and employer businesses discussed earlier. We also discuss our results relative to alternative assumptions about the differences between nonemployer and employer businesses.

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 and aggregate exogenous productivity A 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 in each period t by incurring an entry cost equal to $c_{E,t} \cdot Y_t/L_t$, where $c_{E,t}$ is a relative entry cost parameter and Y_t/L_t is aggregate output per worker in period t. We assume that entry costs scale up with output per capita consistent with the evidence in Bollard et al. (2016) and Bento and Restuccia (2021).[3] We allow relative entry costs $c_{E,t}$ to change over time, along with other exogenous variables, in order to match the evolution of the net entry rate of firms as we explain below. We assume entrants in each period draw their initial productivity z_E from a constant cumulative distribution function $G(z_E)$, and learn their productivity after entry.[4] Each period, conditional on firm survival, firm-level productivity grows exogenously for all firms. We denote by $g_t(z, a)$ the growth rate of firm productivity between periods $t-1$ and t for a firm with productivity z and age a in period $t-1$. We denote by $\lambda_t(z, a)$ the probability that a firm with productivity z and age a in period $t-1$ exits before period t. We allow both $g_t(z, a)$ and $\lambda_t(z, a)$ to change exogenously over time. The fraction of age-a firms with productivity z in period t is denoted by $h_t(z|a)$. We denote the fraction of firms with age a by $x_t(a)$. We assume no fixed overhead cost for producers, but in Appendix E we show that incorporating overhead costs does not change the main implications of the model.

Average productivity across all entrants z_E is equal to the expected value of a draw from $G(z_E)$, denoted by $\bar{z}_{ent} = \int_{z_E} z_E \; dG(z_E)$. We denote average productivity across all firms in period t by $\bar{z}_t = \int_a \int_z z \; h_t(z|a)x_t(a) \; dzda$ and the total number firms by N_t. Let $N_{ent,t}$ denote the

[3]We specify entry costs in terms of the final good, rather than labor, in order to provide a straightforward mapping between labor per firm in the model and in the data.

[4]Firm-level total factor productivity is specified as $(Az)^{1-\alpha}$, but we refer to z as firm-level productivity throughout for ease of exposition.

number of entrants in period t. Conditional on $N_{t\ldots\infty}$ and $N_{ent,t\ldots\infty}$, average productivity $\bar{z}$ evolves according to the following law of motion;

$$\bar{z}_{t+1}N_{t+1} = N_{ent,t+1} \cdot \bar{z}_{ent} + N_t \int_a \int_z z \cdot (1-\lambda_{t+1}(z,a))(1+g_{t+1}(z,a))h_t(z|a)x_t(a)\ dzda, \quad (2)$$

or

$$\bar{z}_{t+1} = \frac{N_{ent,t+1}}{N_{t+1}} \cdot \bar{z}_{ent} + \frac{N_t}{N_{t+1}}\bar{z}_t \cdot (1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1}), \quad (3)$$

where $(1-\bar{\lambda}_{t+1}) \equiv \int_a \int_z (1-\lambda_{t+1}(z,a))h_t(z|a)x_t(a)\ dzda$ denotes the average survival rate from periods t to $t+1$ for all producers existing in period t, and $\bar{g}_{t+1}$, the 'average' growth rate of firm-level productivity from t to $t+1$ conditional on survival, satisfies the following identity;

$$\bar{z}_t \cdot (1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1}) \equiv \int_a \int_z z \cdot (1-\lambda_{t+1}(z,a))(1+g_{t+1}(z,a))h_t(z|a)x_t(a)\ dzda. \quad (4)$$

We denote employment (equal to population) in each period by L_t, and we abstract from household choices by assuming an exogenous time-varying real interest rate R_t. We assume that all firms know with certainty the evolution over time of L_t, $c_{E,t}$, A_t, and R_t. To accommodate the availability of data, we assume all potential entrants believe exit rates and firm-productivity growth rates $\lambda_t(z,a)$ and $g_t(z,a)$ are equal to their across-firm averages, $\bar{\lambda}_t$ and $\bar{g}_t$. Potential entrants know with certainty the evolution over time of $\bar{\lambda}_t$ and $\bar{g}_t$, and hence rationally anticipate future N_t and $\bar{z}_t$.[5] Given that our data ends in 2014, we describe below the additional assumptions we make about firms' beliefs about how the economy evolves after 2014.

4.2 Equilibrium

A *competitive equilibrium* is defined by sequences of wage rates $\{w_t\}_{t=0}^{\infty}$, firm-level functions for labor demand $\{\ell_t(z)\}_{t=0}^{\infty}$, output $\{y_t(z)\}_{t=0}^{\infty}$, and current period profits $\{\pi_t(z)\}_{t=0}^{\infty}$; number

[5]The law of motion for $\bar{z}$, as characterized in equation (3), shows that knowledge of future $\bar{\lambda}_{t+1}$ and $\bar{g}_{t+1}$ are enough to anticipate the evolution of $\bar{z}_t$, given current $\bar{z}_t$, current N_t, and future N_{t+1}, noting that the future number of entrants $N_{ent,t+1}$ is a simple function of N_t, N_{t+1}, and $\bar{\lambda}_{t+1}$.

of entrants $\{N_{ent,t}\}_{t=0}^{\infty}$, and number of firms $\{N_t\}_{t=0}^{\infty}$, given sequences of exogenous relative entry cost parameter $\{c_{E,t}\}_{t=0}^{\infty}$, labor supply $\{L_t\}_{t=0}^{\infty}$, real interest rate $\{R_{t+1}\}_{t=0}^{\infty}$, aggregate productivity parameter $\{A_t\}_{t=0}^{\infty}$; exit rates $\{\lambda_t(z,a)\}_{t=0}^{\infty}$, productivity growth rates $\{g_t(z,a)\}_{t=0}^{\infty}$, cross-sectional distributions $\{h_t(z|a)\}_{t=0}^{\infty}$ and $\{x_t(a)\}_{t=0}^{\infty}$; time-invariant distribution $G(z_E)$; and initial conditions N_{-1}, $h_{-1}(z|a)$, and $x_{-1}(a)$; such that:

(i) In each period, given w_t and A_t, firms choose $\ell_t(z)$ to maximize $\pi_t(z)$, implying $y_t(z)$.

(ii) In each period, free entry ensures the expected present value of lifetime profits for an entrant is less than or equal to the current entry cost, a condition that holds with equality in a period with positive entry;

$$c_{E,t} \cdot Y_t/L_t \geq \pi_t(\bar{z}_{ent}) + \sum_{t'=t+1}^{\infty} \pi_{t'}(\tilde{z}_{t'}) \cdot \prod_{k=t+1}^{t'} \frac{1-\bar{\lambda}_k}{1+R_k}, \tag{5}$$

where $\tilde{z}_{t+1} = \bar{z}_{ent} \cdot (1+\bar{g}_{t+1})$, $\tilde{z}_{t+2} = \tilde{z}_{t+1} \cdot (1+\bar{g}_{t+2})$,..., R_k is the interest rate from period $k-1$ to k, and aggregate output is $Y_t = N_t \int_a \int_z y_t(z) h_t(z|a) x_t(a) dz da$.[6]

(iii) The law of motion for $\bar{z}$, as characterized by equations (2) and (3), holds for all $t \geq -1$, and the law of motion for the number of firms satisfies,

$$N_t = N_{ent,t} + N_{t-1}(1-\bar{\lambda}_t), \qquad t = 0, ..., \infty. \tag{6}$$

(iv) In each period, the labor market clears: the exogenous aggregate labor supply L_t is equal to the aggregate quantity of labor demanded by firms, $L_t = N_t \int_a \int_z \ell_t(z) h_t(z|a) x_t(a) dz da$.

To solve for the competitive equilibrium, we start with the labor decision. Producers in each period choose labor to maximize operating profits, resulting in the following optimal demand

[6] The presence of average productivity of entrants $\bar{z}_{ent}$ inside the profit function exploits the fact that profit is proportional to z, as characterized in equation (9), so that the expected value of $\pi(z_E)$ is equivalent to $\pi(\bar{z}_{ent})$.

for labor, output, and operating profits, expressed as functions of z;

$$\ell_t(z) = A_t z \left(\frac{\alpha}{w_t}\right)^{\frac{1}{1-\alpha}}, \tag{7}$$

$$y_t(z) = A_t z \left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}}, \tag{8}$$

$$\pi_t(z) = A_t z \left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}} (1-\alpha). \tag{9}$$

Note that profit maximization implies a direct mapping from relative employment (7) to relative revenue (8) to relative productivity z. We use these relationships later both to overcome data limitations and to infer relative productivity when we infer how average productivity changes over time in the data.

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

$$L_t = N_t \cdot A_t \left(\frac{\alpha}{w_t}\right)^{\frac{1}{1-\alpha}} \bar{z}_t,$$

where $\bar{z}$, average firm-level productivity across all producers, is defined as in equations (2) and (3) . The wage can therefore be expressed as a function of $\bar{z}$ and the number of firms per worker (N/L) in each period;

$$w_t = \alpha (A_t \bar{z}_t)^{1-\alpha} \left(\frac{N_t}{L_t}\right)^{1-\alpha}. \tag{10}$$

Using equation (10) into equation (9), profit in each period can be expressed as:

$$\pi_t(z) = (1-\alpha) \frac{A_t^{1-\alpha} z}{\bar{z}_t^{\alpha}} \left(\frac{L_t}{N_t}\right)^{\alpha}. \tag{11}$$

Using equations (8) and (10), aggregate output per worker can be expressed as a function of $\bar{z}$ and firms per worker in each period:

$$\frac{Y_t}{L_t} = \frac{N_t}{L_t} \cdot A_t \left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}} \bar{z}_t = (A_t \bar{z}_t)^{1-\alpha} \left(\frac{N_t}{L_t}\right)^{1-\alpha}. \tag{12}$$

Note that aggregate output per capita Y_t/L_t in each period as characterized in equation (12) is only a function of contemporary variables A_t, N_t/L_t, and $\bar{z}_t$. As we show below, the competitive equilibrium values of both N_t/L_t and $\bar{z}_t$ in each period can be derived from the entire sequences of exogenous variables, including relative entry costs. In turn, the sequences of exogenous variables can be appropriately selected to match the data on these endogenous variables. As a result, the static competitive equilibrium conditions derived above allow us to infer the proximate impact on aggregate productivity from changes in firm dynamics. Given data on Y_t/L_t and N_t/L_t, and assuming we can obtain a measure of $\bar{z}_t$ which we also show below, exogenous aggregate productivity A can be inferred as a residual from equation (12). We can then use the expression in equation (12) to decompose changes in output per capita into those arising due to firm dynamics (changes in N_t/L_t and $\bar{z}_t$) and exogenous forces unrelated to firm dynamics (changes in A_t and L_t). We also emphasize that all variables of interest up to this point are simple expressions of $\bar{z}_t$, a measure of average productivity across producers, summarizing all the firm-specific dynamic variables $\lambda_t(z,a)$ and $g_t(z,a)$, as well as cross-firm distributions $h_t(z|a)$ and $x_t(a)$.

We now simplify the free entry condition in equation (5) as follows. First, we note that in our data there is positive entry in every period and hence the entry condition holds with equality. Second, we exploit the linearity of profits with respect to firm-level productivity in equation (11) and our assumption that potential entrants believe exit and productivity growth rates are common across firms in each period, an assumption necessary due to the lack of detailed data by age and size; to write current entry costs as follows:

$$c_{E,t} \cdot (Y_t/L_t) = (1-\alpha)\bar{z}_{ent}\left[A_t^{1-\alpha}\left(\frac{L_t/N_t}{\bar{z}_t}\right)^{\alpha} + A_{t+1}^{1-\alpha}\frac{(1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1})}{1+R_{t+1}}\left(\frac{L_{t+1}/N_{t+1}}{\bar{z}_{t+1}}\right)^{\alpha}\right.$$

$$\left. + A_{t+2}^{1-\alpha}\frac{(1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1})}{(1+R_{t+1})}\frac{(1-\bar{\lambda}_{t+2})(1+\bar{g}_{t+2})}{(1+R_{t+2})}\left(\frac{L_{t+2}/N_{t+2}}{\bar{z}_{t+2}}\right)^{\alpha} + ...\right].$$

Because the free entry condition above holds with equality in arbitrary periods t and $t+1$, entry costs in period t can be written as profits in period t plus appropriately discounted entry costs in period $t+1$:

$$c_{E,t} \cdot (Y_t/L_t) = (1-\alpha)\bar{z}_{ent} A_t^{1-\alpha} \left(\frac{L_t/N_t}{\bar{z}_t}\right)^{\alpha} + \frac{(1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1})}{(1+R_{t+1})} c_{E,t+1} \cdot (Y_{t+1}/L_{t+1}) \, .$$

Further, using equation (12), the above expression can be simplified as:

$$c_{E,t} = (1-\alpha)\left(\frac{L_t}{N_t}\right)\frac{\bar{z}_{ent}}{\bar{z}_t} + c_{E,t+1}\frac{(1-\bar{\lambda}_{t+1})(1+\bar{g}_{t+1})}{1+R_{t+1}}\left(\frac{Y_{t+1}/L_{t+1}}{Y_t/L_t}\right). \tag{13}$$

Equation (13) characterizes the number of firms (and therefore the number of entrants, given equation 6) in every period $t = 0, ..., \infty$. Combined with equations (3) and (12), the number of firms per worker in every period depends on the entire sequences of exogenous variables. More specifically, equation (3) which holds from $t = -1, ..., \infty$ determines the evolution of $\bar{z}_t$ as a function of sequences for $\bar{\lambda}_t$ and $\bar{g}_t$, conditional on the sequence of $N_{ent,t}$, the initial value of N_{-1} and the initial $\bar{z}_{-1}$ which can be obtained from the initial distributions $h_{-1}(z|a)$ and $x_{-1}(a)$; equation (12) characterizes Y_t/L_t in each period as a function of contemporary A_t, N_t, $\bar{z}_t$, and L_t; and equation (6) maps sequences of N_t into $N_{ent,t}$ given sequences of $\bar{\lambda}_t$ and initial condition N_{-1}.

The above discussion suggests we require knowledge of the entire sequences of all exogenous variables in order to solve for N/L and $\bar{z}$ in any period. We focus on the period between 1982 and 2014. To deal with the lack of data after 2014, we note that equation (13) indicates that to determine the number of firms (and therefore entrants) per worker in a given period t, it is sufficient information about outcomes after period t for potential entrants to know the growth in output per worker, the exit rate, the growth in firm productivity, and the interest rate from period t to $t+1$. In this context, in order to use equations (3) and (13) to solve for N_t/L_t and $\bar{z}_t$ in each period we consider given the available data, we further assume all potential entrants believe relative entry costs c_E, the interest rate R, the average exit rate $\bar{\lambda}$, the average firm productivity growth rate $\bar{g}$, and the growth rate of aggregate output per worker Y/L in 2015 are the same as in 2014, and that there is positive entry in 2015. With measures of c_E, $\bar{\lambda}$, $\bar{g}$,

R, A, and L from 1982 to 2014, as well as $\bar{z}$ in 1981 and $\bar{z}_{ent}$, under the above assumptions we can solve for N and $\bar{z}$ in each period we consider. Conversely, with measures of N/L, $\bar{z}$, $\bar{\lambda}$, $\bar{g}$, R, A, and L from 1982 to 2014, we can infer values for c_E in each period. We now turn to characterizing average productivity $\bar{z}$.

4.3 Inferring Average Productivity

The model provides a mapping from relative employment size to relative productivity, given by equation (7). Denoting the average employment size of all firms by $\bar{\ell}$ and entrants by $\bar{\ell}_{ent}$, average productivity of all firms relative to entrants in each year is;

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

We emphasize that equation (8) implies that the same mapping of relative productivity to relative employment applies between productivity z and output (revenue) y, hence, the above expressions hold when substituting revenue for labor. This is relevant because for nonemployers we do not have data on employment, only revenue. If we have a value for $\bar{z}_{ent}$, we can therefore infer $\bar{z}$ using only differences in average firm size across all firms and across entrants. Note that given relative size data, detailed data on $g_t(z, a)$ and $\lambda_t(z, a)$ is not required to infer average productivity, we only require the size of all firms relative to entrants.

We obtain a value for $\bar{z}_{ent}$, assumed time-invariant and equal to the average draw of z_E from $G(z_E)$, by exploiting the availability of data from the SBO 2012 reporting moments from the revenue size distribution of firms by age. We assume $G(z_E)$ can be described by a Pareto distribution, and denote the shape parameter by ξ. The lower bound of the distribution is a free parameter in this framework (not separately identifiable from A), so we normalize it to 1. To obtain a value for ξ, we target the ratio of revenue for the 55th percentile entrant to average revenue across all entrants, equal to 0.1485.[7] The required value for ξ to generate this

[7] This data is only reported for 2012. We choose the 55th percentile because it is the closest reported percentile to the median. Average revenue is $67,359, while revenue for the 55th percentile firm is $10,000.

ratio is 1.076. This implies a value for $\bar{z}_{ent}$ equal to $\frac{\xi}{\xi-1} = 14.2$. Parameterizing $G(z_E)$ in this way gives us a way to infer how average productivity across subgroups of entrants changes over time, using the properties of a Pareto distribution combined with the available data. We detail how we do this below.

To measure average firm-level productivity $\bar{z}$ in each period, given the available data, it is useful to decompose $\bar{z}$ and $\bar{z}_{ent}$ as follows;

$$\bar{z} = \frac{\bar{z}_{ent}^{non} N_{ent}^{non} + \bar{z}_{inc}^{non} N_{inc}^{non} + \bar{z}_{ent}^{emp} N_{ent}^{emp} + \bar{z}_{inc}^{emp} N_{inc}^{emp}}{N}, \tag{14}$$

$$\bar{z}_{ent} = \frac{\bar{z}_{ent}^{non} N_{ent}^{non} + \bar{z}_{ent}^{emp} N_{ent}^{emp}}{N_{ent}} \tag{15}$$

where $\bar{z}$ is an average of the productivities of nonemployer and employer entrants, and nonemployer and employer incumbents, weighted by the fraction of all firms in each category. Average productivity of entrants $\bar{z}_{ent}$ is decomposed in an analogous way. This decomposition illustrates that $\bar{z}$ can be calculated given values for each of the eight variables on the right-hand side: four variables on the number of entrant and incumbent firms for employers and nonemployers; and four variables related to the average idiosyncratic productivity for each type of firm. As a result, with separate measures of labor ℓ (or revenue) for entrant and incumbent nonemployers and employers, in addition to measures of the number of firms of each type, equations (14) and (15) can be used to infer changes in average productivity relevant for equation (12).

While the BDS data provide information for entrant and incumbent employer firms, the NES data do not distinguish between entrant and incumbent nonemployers. Similarly, while the revenue share of nonemployers (for some years) is reported, the data do not distinguish between entrants and incumbents except in 2012. To separate nonemployer entrants and incumbents, we use additional data and assumptions as we specify below. We proceed in three steps.

First, to infer the number of nonemployer entrants N_{ent}^{non}, we use information on exit rates. In particular, we use a survival rate of 85% for nonemployers in 1997 from Davis et al. (2009)

With a Pareto distribution, the ratio of 55th percentile over average is equal to $\frac{\xi-1}{\xi}(1-0.55)^{\frac{-1}{\xi}}$.

and given the lack of exit data for nonemployers over time, we assume that the survival rate of nonemployers changes over time in proportion to the survival rate for employers, which we take from the BDS data. Hence, we impute the survival rate of nonemployers as follows:

$$(1 - \lambda_t^{non}) = 0.85 \cdot \frac{(1 - \lambda_t^{emp})}{(1 - \lambda_{1997}^{emp})}.$$

Given that the exit rate for employers from BDS data decreases from 10% in 1982 to 7.5% in 2014, the implied exit rate for nonemployers decreases from 15.8% to 13.5%. Using these exit rates and data on the total number of nonemployers each year N_t^{non}, we calculate the number of nonemployer entrants as:

$$N_{ent,t}^{non} = N_t^{non} - (1 - \lambda_{t-1}^{non}) \cdot N_{t-1}^{non}. \tag{16}$$

With the total number of nonemployers in each year, the number of incumbent nonemployers is given by $N_{inc,t}^{non} = N_t^{non} - N_{ent,t}^{non}$. The number of incumbent employer firms N_{inc}^{emp} is simply calculated from data on the total number of employer firms N^{emp} and the number of employer entrants N_{ent}^{emp}.

Note that we abstract from nonemployer to employer transitions here since the data does not differentiate between new employers and transitioning nonemployers. As a result, we potentially overstate the number of employer entrants, but the bias is unlikely substantial given the small number of transitioning nonemployers to employers in a given year documented in Davis et al. (2009). It is also possible that nonemployer to employer transitions are increasing over time which could be more relevant, unfortunately there is not enough data to assess this possibility over the time period we consider.

The above analysis implies that as a fraction of the total number of firms, the share of nonemployer incumbents increases from 57% in 1982 to 68.9% in 2014, whereas the share of employer incumbents decreases from 23.7% to 16.1%, and the shares of nonemployer and employer entrants decrease from 15.9% to 13.6% and from 3.4% to 1.4%.

Second, to infer average idiosyncratic productivity of both employer and nonemployer entrants in each period, note that the Pareto distribution of entrant z_E implies;

$$\bar{z}_{ent}^{emp} = \frac{\xi}{\xi - 1}\left(\frac{N_{ent}}{N_{ent}^{emp}}\right)^{\frac{1}{\xi}}, \tag{17}$$

while the decomposition of entrant productivity implies;

$$\bar{z}_{ent}^{non} = \frac{\bar{z}_{ent} N_{ent} - \bar{z}_{ent}^{emp} N_{ent}^{emp}}{N_{ent}^{non}}. \tag{18}$$

Third, to infer average productivity of employer and nonemployer incumbents, we use labor and revenue data. Using BDS data on employment at employer firms for entrants and incumbents, we calculate employer incumbent productivity as;

$$\bar{z}_{inc}^{emp} = \bar{z}_{ent}^{emp} \cdot \left(\frac{\bar{\ell}_{inc}^{emp}}{\bar{\ell}_{ent}^{emp}}\right). \tag{19}$$

We do not have data on employment at nonemployer firms (unpaid or owner-provided). To infer average idiosyncratic productivity of incumbent nonemployers $\bar{z}_{inc}^{non}$, we therefore use data on the revenue share of nonemployers from the SBO, which the model assumes is equal to their employment share. Denote the revenue share of nonemployers as $Rev^{non} \equiv Y^{non}/Y$. The model implies the following relationship between the revenue share of nonemployers and average productivity $\bar{z}$;

$$Rev^{non} = \frac{\bar{z}_{ent}^{non} N_{ent}^{non} + \bar{z}_{inc}^{non} N_{inc}^{non}}{\bar{z}_{ent}^{non} N_{ent}^{non} + \bar{z}_{inc}^{non} N_{inc}^{non} + \bar{z}_{ent}^{emp} N_{ent}^{emp} + \bar{z}_{inc}^{emp} N_{inc}^{emp}},$$

which we use to infer the average idiosyncratic productivity of nonemployer incumbents;

$$\bar{z}_{inc}^{non} = \frac{Rev^{non} \cdot (\bar{z}_{ent}^{emp} N_{ent}^{emp} + \bar{z}_{inc}^{emp} N_{inc}^{emp}) - (1 - Rev^{non}) \cdot z_{ent}^{non} \cdot N_{ent}^{non}}{(1 - Rev^{non}) \cdot N_{inc}^{non}}. \tag{20}$$

Note that we have data for the revenue share of nonemployers in 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, and calculate the corresponding revenue share. We then calculate the average annual growth rate of this measure between 1987 and 2012, and use this growth rate to impute values before 1987 and after 2012. Our imputed revenue share of nonemployers increases from 2.71% in 1982 to 3.11% in 2014, and (as in the data) changes from 2.80% in 1987 to 3.07% in 2012.

To summarize, equations (17) through (20) identify the values for $\bar{z}_{ent}^{emp}$, $\bar{z}_{ent}^{non}$, $\bar{z}_{inc}^{emp}$, and $\bar{z}_{inc}^{non}$, which provide all the additional information required to calculate average productivity $\bar{z}$ using equation (14). Keeping in mind the value for $\bar{z}_{ent} = 14.2$, this analysis implies that between 1982 to 2014, $\bar{z}_{ent}^{non}$ increases from 2.0 to 2.4, $\bar{z}_{inc}^{non}$ from 2.4 to 3.6, $\bar{z}_{ent}^{emp}$ from 72.1 to 129.0, and $\bar{z}_{inc}^{emp}$ from 242.3 to 530.9. Combined with our measures of each type of firm, equation (14) therefore implies an increase in overall average idiosyncratic productivity $\bar{z}$ from 61.5 to 90.1, a cumulative increase of 46.6% from 1982 to 2014. And note that because we target both the numbers of nonemployers and the share of revenue from nonemployers, we exactly match this data by construction.

4.4 Implications

We now discuss the quantitative implications of the change in the total number of firms for aggregate productivity and other variables of interest.

Aggregate productivity. We define aggregate total factor productivity (TFP) in the model analogous to how it is measured in the data, aggregate output per unit of aggregate composite variable inputs, which is a constant-returns-to-scale function of aggregate variable inputs. As such, we refer to aggregate TFP in the model as measured TFP. We do this in order to make direct comparisons of how changes in firm net entry contribute to changes in measured TFP in the U.S. data. In the model, measured aggregate TFP is the same as aggregate output per worker Y/L.[8]

We note that while measured aggregate TFP (equation 12) depends on the common productivity term A, the number of firms per worker does not (though it does depend on A growth). This implies that in our quantitative analysis, we can treat A as a free parameter that captures all changes in measured TFP in the data that are not otherwise accounted for by the forces captured by the model such as the change in the number of firms per worker. As a result, we focus below on the change in TFP associated with the change in the number of firms per worker, as well as with the change in average firm-level productivity $\bar{z}$.

Using our measures for $\bar{z}_t/\bar{z}_{ent,t}$ and data for N_t/L_t, we calculate implied aggregate TFP in each period from the last two terms in equation (12). Figure 3a reports the evolution of implied aggregate TFP due to changes in $\bar{z}$ and N/L. Figure 3b compares the combined change in TFP due to changes in firm dynamics (both $\bar{z}$ and N/L) with observed TFP in the data.[9] We also show the implied evolution in A (calculated as a residual) for completeness.

The TFP impact of changes in average firm-level productivity ($\bar{z}$) in Figure 3a (normalized to 1 in 1982), is generally flat from 1982 to 1997, then increases by an average 0.39% per year from 1997 to 2014. As a result, the cumulative increase in aggregate TFP due to average productivity is 8.0% from 1982 to 2014. The corresponding impact of the increase in firms per worker is a cumulative 7.1%, implying an average annualized growth rate in implied TFP of 0.22%. Taken together, the evolution in firm dynamics over time implies a cumulative 15.6% increase in aggregate TFP from 1982 to 2014, as illustrated in Figure 3b. The cumulative growth in TFP observed in the data between 1982 and 2014 amounts to 32%, hence the impact of firm dynamics accounts for about half of the actual increase in TFP.

Aggregate TFP growth implied by the model is decomposed into two parts in Figure 3a. First, the increase in the number of firms per person by itself implies cumulative growth in TFP of

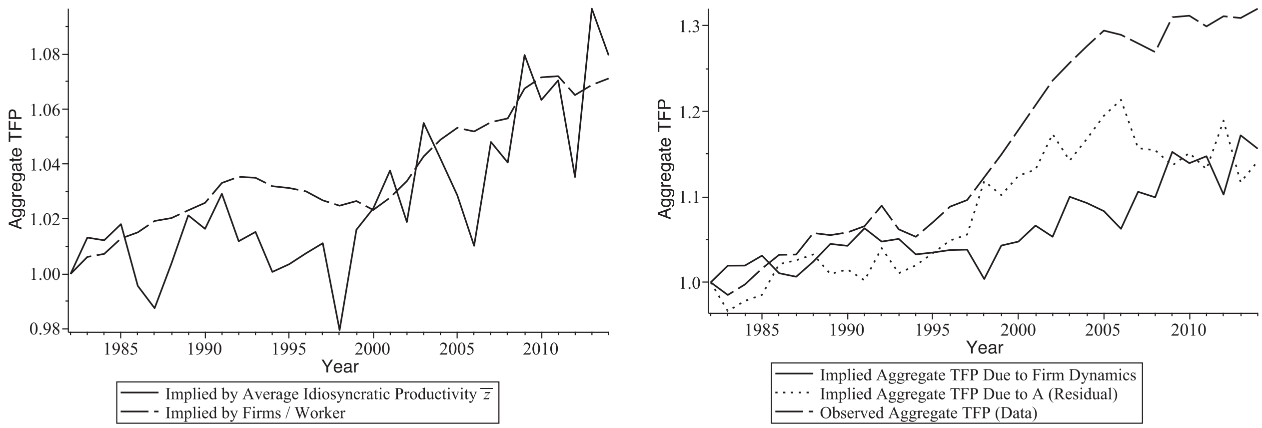
[8]In a frictionless capital accumulation environment there is a simple mapping from aggregate TFP to aggregate labor productivity. Nevertheless, we focus on the comparison of our model outcomes with measured aggregate TFP in the data since this moment abstracts from any potential distortion to capital accumulation in the data.

[9]Data for aggregate TFP are from Fernald (2012).

Figure 3: Aggregate TFP

(a) Decomposition: Average Productivity and Firms/Worker

(b) Decomposition: Firm Dynamics and Exogenous TFP



Notes: Panel (a): The solid line reports changes over time in aggregate TFP implied solely by changes in average idiosyncratic productivity $\bar{z}$. The dashed line reports the same implied by changes in firms per worker. Panel (b): The solid line reports total changes over time in aggregate TFP implied by changes in $\bar{z}$ and firms per worker. The dotted line reports how TFP changes due to changes in exogenous aggregate productivity A. The dashed line reports how observed aggregate TFP changes in the data, from Fernald (2012).

7.1% from 1982 to 2014.[10] This growth in productivity is the result of smaller firms on average and hence a higher marginal (and average) product of labor, driven by decreasing returns to scale in each firm's production function. Second, the shift in the firm size distribution suggests that incumbents have become more productive than entrants over time, due to a combination of lower exit rates and higher productivity growth over the life of the average firm. This shift suggests average productivity across all firms has increased by an additional 8.0% since 1982. We emphasize that our analysis allows for any underlying exit and growth dynamics by size and age among incumbent employers. What ultimately matters for our calculation is the average $\bar{z}$ among incumbents, relative to entrants (all relative to 1982), regardless of which incumbent employers are growing more or less, or surviving with higher or lower probabilities.

We contrast our results with those implied by firm dynamics when abstracting from nonemployers. We derive the implied growth in aggregate TFP using only employer data from BDS, a straightforward assessment given that the numbers of entrants and incumbents, as well as

[10] When calculating firms per worker in Section 3 we use CPS data for our measure of aggregate employment. In Appendix B we recalculate firms per worker using a measure of aggregate employment inferred from the BDS data for robustness.

their relative average size, are directly reported in the data. During the same time period from 1982 to 2014, we find that the implied aggregate TFP due to firm dynamics increases by only 2.1%, compared to 15.6% when accounting for nonemployers. The lower aggregate TFP growth associated with employers arises mostly due to a decline in employers per worker, generating a cumulative -4.8% drop in aggregate TFP (7.1% increase with nonemployers), which is offset by an increase in aggregate TFP of 6.9% due to rising average productivity $\bar{z}$ (8% increase with nonemployers). See appendix F for more details. Interestingly, accounting for nonemployers implies a larger TFP growth mostly due to the increase in firms per worker, but also due to an increase in resource reallocation driving higher average firm-level productivity. These contrasting results between firm dynamics characterized with and without nonemployers highlight the quantitative importance of a more comprehensive measure of business dynamism in the U.S. data.

We also note that a 10-year moving average of our measures of implied TFP growth from firm dynamics features a much higher correlation with productivity growth in the data (21.3%) compared to a strongly negative corresponding correlation using only employer data as above (-42.7%), and compared to the literature's findings based on only employer data. This again suggests that net firm entry when including nonemployers in fact plays an important quantitative role in driving trends in aggregate productivity. While our analysis is not designed to account for the short- and medium-run movements in TFP that surely depend on many factors not included in the model, we emphasize that our framework significantly improves on the strong negative correlation between observed TFP growth and that implied by a model with employer firms as highlighted by Decker et al. (2016) and Li (2017).

Since the model implies that the shares of employment and revenue of nonemployers are the same, we use revenue shares of nonemployers to bypass the lack of data on employment when calculating $\bar{z}$ above. This implicitly assumes that revenue per unit of labor for nonemployers is the same as revenue per labor of all firms. Is this a reasonable assumption? Presumably, revenue per labor of nonemployers is closer to small employers than to all employers. We look at disaggregated employer data from the U.S. Census' Statistics of U.S. Businesses for 2012 and find that revenue per employee in the smallest employers (less than 5 employees) is equal to or greater than revenue per employee for all employers in 70 percent of the 2-digit industries (14 industries out of 20). It is also not clear there are systematic patterns of change over time. For instance, across all industries, revenue per employee across small employers is 13% lower than that across all employers in 2012 but about 8% higher in 1997. Nevertheless, we assess the potential relevance of our assumption using the revenue per employee of small employers from 2012. To do this, we redo our calculations of $\bar{z}$ in each year assuming average employer productivities $\bar{z}^{emp}_{inc}$ and $\bar{z}^{emp}_{ent}$ are 14.9% higher than we previously calculated.[11] We find that the productivity effect of this adjustment is very small, an implied TFP decrease between 1982 and 2014 due to firm dynamics of 15.63%, compared with 15.64% in our benchmark analysis. Why is the implied change in TFP almost identical? Increasing the implied productivity of employers increases $\bar{z}$ in all years, which need not affect the change over time. But while $\bar{z}^{emp} \cdot N^{emp}$ increases by 14.9% in every year due to our adjustment, $\bar{z}^{emp} \cdot N^{emp}$ as a fraction of $\bar{z} \cdot N_{all}$ decreases slightly over time in our benchmark analysis from 97.3% to 96.9%. As a result, on net, the adjustment increases $\bar{z}$ slightly less in 2014 than in 1982.

Motivated by the comparison between nonemployers and employers discussed in Section 2, our approach is to treat nonemployers the same as employers but simply less productive. A plausible alternative is to consider a setting where all entrants pay an entry cost, but then make a costly decision about whether to be an employer which results in a productivity improvement relative to nonemployers. This setup is equivalent to one where firms choosing to be nonemployers face some exogenous drop in effective productivity. We emphasize that there is not sufficient data to discipline this extension of the model, mainly because it would be difficult to separately identify the underlying productivity of employers and the increase in productivity they obtain from becoming employers. However, we note that we can calculate $\bar{z}$ for any given increase in productivity associated with being an employer, much as we do above when considering

[11] If nonemployers are 13% less productive relative to the average than we previously assumed, then employer productivity is $1/0.87 = 1.149$ times higher than we previously calculated.

differences in revenue per unit of labor. We do this for a large increase in z for all employers, equivalent to an arbitrary 50% increase in productivity. We find that the impact on aggregate TFP from firm dynamics drops from a 15.64% increase in our baseline to 15.58% in this case. This result suggests that extending the model to consider employer status as conferring benefits to firms, is unlikely to change substantially the overall impact on aggregate TFP from the change in the number of firms.

In Section 3 we discuss a concern with nonemployer counts in that they may contain persons that are effectively employees but register their earnings as business income for tax purposes. Figure 2 shows that when we exclude sole-proprietor nonemployers from our firm counts, the total number of firms per worker still rises by 7% from 1997 to 2014, compared to a 24% increase in our baseline data over the same period. As a robustness, we recalculate the growth in implied aggregate TFP due to changes in the number of firms per worker and average productivity $\bar{z}$ when sole-proprietor nonemployers are excluded. We do not have the data to re-estimate a value for the Pareto parameter related to $G(z_E)$, hence in this calculation, we assume the same value of $\xi = 1.076$ as in our baseline analysis. We recalculate average productivity $\bar{z}$ for 1997 and 2014 following Section 4.3. In our baseline, aggregate TFP grows by a cumulative 11.4% from 1997 to 2014 due to the evolution in firm dynamics. Without sole-proprietor nonemployers, the cumulative growth in aggregate TFP is 5.9%. To the extent that not all the sole-proprietors nonemployers counts are effectively employees in employer firms, this lower TFP growth can be interpreted as a lower bound of the change in the total number of firms. More importantly, this lower implied growth abstracting from sole-proprietors is still much higher than the growth implied when abstracting from nonemployers entirely.[12]

Entry costs. Using the free entry condition as characterized by equation (13) along with our measures of how $\bar{z}_{ent}/\bar{z}$, N/L, $\bar{\lambda}$, $\bar{g}$, R, and Y/L evolve over time, we can infer values for c_E for each year t. We illustrate our results in Figure 4.

[12] Figure F.3 Panel (b) in Appendix F shows that aggregate TFP implied by the model when nonemployers are completely excluded declines slightly from 1997 to 2014.

Figure 4: Implied Entry Costs relative to GDP

Notes: All values are reported relative to c_E in 1982.

While entry costs (relative to aggregate output per worker) rise in the 1990s, rationalizing the drop in firms per worker, the implied persistent drop after 1998 is dramatic, resulting in entry costs that are 66 percent lower in 2014 than in 1982. The drop in relative entry costs over time are even more dramatic than the increase in the number of firms per worker. This suggests the expected discounted profits of entrants dropped quite a bit over time. The combination of lower average exit rates and higher average productivity growth rates over time led to higher average profitability for incumbents. Although entrants could expect to reap the same benefits later in their life cycle, discounting means the negative impact of higher-productivity competition at entry more than offsets the benefit of higher future expected profits. While lower relative entry costs would be required to rationalize the increase in firms per worker even if the average productivity of incumbents had stayed constant (relative to entrants), observed changes in firm dynamics imply an even larger drop in entry costs relative to output per worker.

The implied change in entry costs depends on our assumption that α, the average share across firms, is constant over time. Allowing this parameter to change could therefore lead to a different path for entry costs. Equation (13) suggests that a decrease in α increases the number of firms, essentially by increasing firm profit margins. Note that such a decrease in α would lower each firm's labor expenditure relative to revenue. But Kehrig and Vincent (2021) report a small *increase* in this ratio over the time period we consider, almost 2% higher. This suggests that changes in α are not driving the dramatic rise in the number of firms per worker we document.

Is a decline of 66% in relative entry costs plausible as a source of the increase in the number of firms per worker? Keeping in mind the lack of direct measures for the costs of starting and operating a firm, we discuss two pieces of supporting evidence: a direct qualitative measure and an indirect quantitative measure. First, the OECD (Organisation for Economic Co-operation and Development) constructs an index of how restrictive regulatory barriers to entrepreneurship are for a number of countries, using objective data on *de jure* laws and regulations (Koske et al., 2015). They report a consistent decline in barriers from 1998 to 2013 in the United States, consistent with Figure 4. Second, Bollard et al. (2016) use a framework similar to ours to infer how entry costs change over time. Acknowledging lack of data on nonemployers, they infer changes in entry costs using firm-level data on value-added over time. They note that the present discounted value of lifetime value-added for entrants should be proportional to entry costs, as in our framework. Their analysis suggests that total entry costs have decreased by about 50% from 1982 to 2012 (from Figure 1 in Bollard et al., 2016), relative to aggregate output per worker. This is very similar to the 66% decline we infer from our data.

The number of nonemployer firms. Our model characterizes nonemployers as firms demanding less than 1 unit of labor with labor supplied by an owner and our analysis exactly matches the number of nonemployer firms in the data over time by construction.[13] Through the lens of our model, the data suggest nonemployers have become a greater share of all firms over time for two reasons. First, the productivity distribution underlying the firm-size distribution has evolved over time, shifting a larger share of economic activity to the most productive firms and away from the least productive. Second, the overall increase in the number of firms per worker implies a lower average size of all firms.

To understand why these two trends result in a larger share of nonemployers, it is useful to

[13]This implies that owners of nonemployer firms use any remaining labor endowment to supply labor to employer firms and earning a competitive wage, consistent with our discussion of the self-employment data in Section 2.

use equations (7) and (10) to characterize how a firm's optimal labor demand in equilibrium depends on both relative productivity and average firm size;

$$\ell = \left(\frac{z}{\bar{z}}\right)\left(\frac{L}{N}\right). \tag{21}$$

If the number of firms per worker increases, demand for labor increases and pushes up the equilibrium wage (equation 10). This results in lower employment for any given level of productivity. And the higher is a firm's productivity relative to the average, the higher its labor demand will be relative to the average. Using equation (21) we can solve for a productivity threshold z_1 below which a firm is counted as a nonemployer;

$$z_1 = \bar{z}\left(\frac{N}{L}\right). \tag{22}$$

The number of nonemployers per worker is the number of firms below the productivity threshold z_1, which is increasing in both the number of firms per worker and average idiosyncratic productivity $\bar{z}$. If the number of firms per worker increases, keeping $\bar{z}$ constant, the fraction of firms that are nonemployers increases as z_1 rises. And if an increase in $\bar{z}$ is accompanied by a change in the shape of the z-distribution, such that a greater fraction of firms fall below z_1, then again the fraction of firms that are nonemployers increases.

How much of the observed rise in the nonemployer share over time is due to each of these factors? To answer this question, in Appendix G we perform two counterfactual exercises in order to isolate the impact on the share of nonemployers from changes in the number of firms per worker and changes in across-firm distribution of relative productivity. The results of these counterfactuals indicate that both the increase in the number of firm per worker and changes in the firm-level productivity distribution are equally important in driving the rise of the share of nonemployers observed in the data between 1982 and 2014.

5 Alternative Drivers of Firm Dynamics

There is a growing literature in macroeconomic development linking lower average firm size (the ratio of workers to the number of firms) and lower aggregate productivity in less developed countries to institutional and policy distortions. In this context, we ask whether the increase in the number of firms, and hence the observed decline in average firm size when including nonemployers in the U.S. economy, may be due to increasing distortions rather than the changes in exit and productivity growth patterns or fixed costs considered in our model, which would affect the contribution of the increase in the number of firms to growth in aggregate TFP. We evaluate the possibility of increased factor misallocation, as well as increased entry barriers for new markets. Using the limited data available, we do not find substantial evidence for these channels.

Increasing misallocation among employer firms. The literature investigating the extent and impact of the misallocation of productive factors across firms shows that 'random' misallocation need not affect the equilibrium number of firms. But if larger and more productive firms are effectively taxed at higher rates (or face tighter constraints) 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, which reduces aggregate TFP. 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 (specifically, how quickly the level of firm-level distortions increase with 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 there is evidence of an increase in systematic misallocation in the U.S. economy over time.

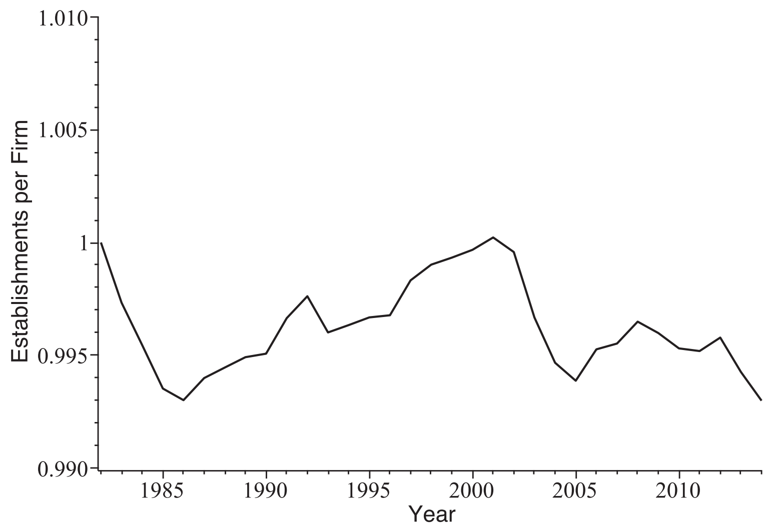
To assess this mechanism, we use publicly-available data from the Economic Census for 74 industries (3-digit NAICS) for the years 2002 and 2012. Although we only have data for 2002 and 2012, this period is still characterized by a substantial increase in the number firms (22%). For each year and in each industry, we have data for 12 employment size bins on the number of firms, wage bill, revenue, and the average number of employees per firm. Note that for manufacturing industries, we use establishments rather than firms and value added instead of revenue. For each year, industry, and employment-size bin, we consider the ratio of revenue to the wage bill (average revenue product) as our measure of the average distortion faced by firms within the bin. Hsieh and Klenow (2009) show that, under certain structural assumptions, profit maximization implies that in the absence of distortions, firms choose labor inputs to equalize the average revenue product of labor across them. If the average revenue product rises with average employment size per firm across bins in an industry-year, we interpret it as evidence that larger firms face larger implicit distortions, and we are interested in finding evidence for this gradient to have increased between 2002 and 2012. This analysis only uses employer firms, since it relies on firms with employees. But if the elasticity of distortions with respect to size is constant across firms within an industry-year, then our elasticity measure is well-identified.

In each year and industry, we regress log average revenue product on a constant and log employees per firm across bins, using the number of firms in each bin as weights. For each year, we obtain an estimate of the industry-specific elasticity of distortions with respect to employment size as our measure of misallocation and its standard error. We then use these estimates to compute the difference in elasticities between 2002 and 2012. We find that on average the change in elasticities over time is neither economically nor statistically significantly different from zero (an average increase of 0.006 with a 95 percent confidence interval of -0.033 to 0.046). This evidence suggests there is no change in misallocation among employer firms. Moreover, only 13.5 percent of the industries, which represent only 5.7 percent of all employer firms, have a significantly positive increase in the elasticity (an increase high enough to fall outside the 95 percent confidence interval). To address the potential issue of unpaid workers not being counted in small firms, we have reproduced these estimates using only firms with at least 5 employees, with the resulting estimated changes in elasticities even closer to zero, reinforcing our findings.

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) 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 5 reports that the number of establishments per firm essentially remained constant from 1982 to 2014, even as the number of firms per worker grew by 41%. 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 admittedly limited and simple evidence presented in this section suggests that various mechanisms encouraging firm entry while lowering aggregate TFP, which may be prevalent in developing countries, are not likely driving the increase in firm net entry or the increase in the share of nonemployer firms in the United States. Future work with more detailed data may be needed to reassess these conclusions.

Figure 5: The Number of Establishments per Firm (Relative to 1982)



Notes: Establishment counts are nonemployers (who we assume have one establishment per firm) plus total establishments owned by employer firms from Business Dynamics Statistics (BDS), U.S. Census Bureau.

6 Conclusions

An important literature documenting a decline in business dynamism in the U.S. economy over the last several decades has focused solely on employer firms. We construct a broader measure of firms that includes nonemployers, firms without paid employees, and find that the total number of firms has diverged dramatically from the number of employer firms over time. In particular, between 1982 and 2014, the total number of firms per worker increased by 41%, whereas during the same period the number of employer firms per worker decline by −8.7%. We interpret this fact, along with the evolution of the employment distribution across firms, through the lens of a standard model of firm dynamics based on Hopenhayn (1992), extended to include nonemployers. We show that accounting for nonemployers drastically changes the implications of firm dynamics for aggregate productivity. Although nonemployers are small relative to employers, the increase in the number of firms per worker has been responsible for around one-quarter of the observed aggregate productivity growth from 1982 to 2014. The accompanying shift in the relative firm size distribution over time has been responsible for another one-quarter of observed growth. These conclusions are in striking contrast to the almost no impact on aggregate productivity implied by the model when abstracting from nonemployer 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 measured TFP growth, casting doubt on the quantitative importance of theories of firm dynamics. Our broader measure of firm net entry, 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, an important question remains as to what accounts for the productivity slowdown in the U.S. economy in recent decades. Our results suggest it is not a decline in firm net entry. One promising recent study instead focuses on the decline in 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. (2021). 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 net 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. In particular, our framework assumes that entering firms operate at their optimal scale, while in practice there may be frictions that affect their operational scale and changes in these frictions could affect aggregate productivity. In particular, if financial constraints have gotten worse for entrants over time relative to other firms, then this could account for some of the increase in the relative employment or revenue of incumbents to entrants. This feature would imply lower productivity for entrants, constituting a drag on aggregate productivity growth. We leave these important explorations for future research.

References

Abraham, Katharine G., John Haltiwanger, Kristin Sandusky, and James Spletzer, "The Rise of the Gig Economy: Fact or Fiction?," *AEA Papers and Proceedings*, 2019, *109*, 357–361.

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.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

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*, 2021, *117*, 220–242.

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.

Guner, N., G. Ventura, and Y. Xu, "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

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.

Kehrig, Matthias and Nicolas Vincent, "The Micro-Level Anatomy of the Labor Share Decline," *Quarterly Journal of Economics*, 2021, *136* (2), 1031–1087.

Koske, Isabell, Isabelle Wanner, Rosamaria Bitetti, and Omar Barbiero, "The 2013 update of the OECD's database on product market regulation: Policy insights for OECD and non-OECD countries," *OECD Economics Department Working Papers*, 2015, *1200.*

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.

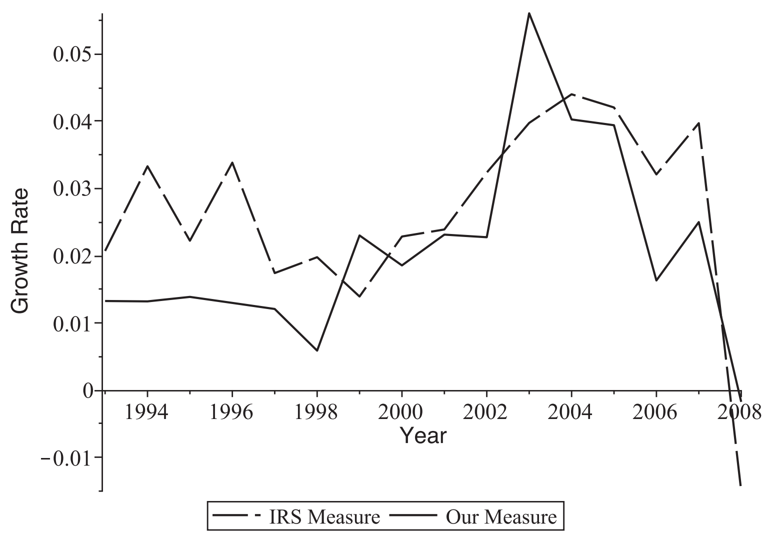
Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration," *NBER Macroeconomics Annual 2020*, 2021, *35*, 115–150.

Appendix

A Data Imputation

For the years between 1981 to 1991, we impute the number of nonemployers 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 A.1 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 A.1: The Growth Rate in the Number of Firms, IRS and Our Measure



Notes: Data from Nonemployer Statistics (NES), Business Dynamics Statistics (BDS), and Statistical Abstract, U.S. Census Bureau.

B Alternative Measures of Aggregate Employment

We assess our main results using alternative measures of aggregate employment.

Accounting for farm workers in the CPS. Farm workers are included in the CPS data we use for aggregate employment, while farms are not included in our firm counts. We do not have data specifically for farm workers, which is why we use the aggregate CPS data in our calculations. But the CPS does report the total number of workers in farming, fishing, and forestry occupations starting in 1983. Although this includes workers in agricultural firms accounted for in our firm data, we can nevertheless check whether removing these employees results in any difference in aggregate labor trends over time. We find it does not. After removing these workers, the overall increase in aggregate employment is exactly the same as for the aggregate CPS data we use in the main body of our paper. As a consequence, using this alternative measure of aggregate employment does not affect any of our results.

Aggregate employment from BDS data. In Section 4 we combine BDS employment data for employers with revenue share data for both employers and nonemployers to infer the share of aggregate employment in nonemployer firms. But we continue to use CPS data as our measure of aggregate employment. It is important to note that these two data sources capture two different concepts of employment – the CPS is measuring workers, while the BDS is measuring formal jobs. A worker can be employed in multiple jobs (formal and informal), while a job may employ either a part-time or full-time (formal) worker. Clearly a better measure of employment across firms for our purposes would be number of hours, but we use BDS data as the best available alternative. For aggregate employment, we believe the CPS measure matches the model most closely. Although it does not account for hours, neither does the BDS data. And the CPS has two important advantages. First, it does not double-count workers with multiple jobs. Second, it accounts for workers not being paid formal wages—owners, family workers, under-the-table workers, among others. Nevertheless, we can recalculate our measure of firms per worker each year and the contribution of its increase to aggregate TFP, using the same strategy to infer employment at nonemployers based on BDS employment and revenue share data, and simply adding this to total BDS employment (ignoring the CPS). When we do this, we find aggregate employment increases by a larger 61% from 1982-2014, compared to 54% in the CPS data. As a result, the number of firms per worker grows by a lower 35%, rather than 41%. This translates into a cumulative increase in aggregate TFP due to the increase in the number of firms per worker of 6.2%, compared to the 7.1% we report in Section 4.4.

C Criteria for Inclusion in Nonemployer Counts

The minimum revenue threshold for inclusion in the NES data has remained constant in nominal terms, at $1,000 for most industries over the time period we examine. This implies that some of the increase in the measured number of nonemployers we documented might be the result of a decreasing inflation-adjusted revenue threshold over time. To examine this potential issue, we use available data for the 2012 Census year, where we have the proportion of nonemployer businesses that earned between $1,000 and $5,000 in revenue, the lowest and narrower category of nonemployer businesses reported. Using the GDP deflator between 1982 and 2012, we calculate that $1,000 in 1982 represent $2,036 in 2012. This implies that about one quarter of the range in the lowest revenue category may be affected. As a result, assuming a uniform distribution of nonemployer businesses in the lowest category, we estimate that one quarter of the businesses in the lowest revenue category were existing businesses in 1982 that were not counted and removing these businesses implies that the growth in the total number of firm between 1982 and 2012 is 31 percent compared with the 37 percent in our data. That is, this adjustment implies that the growth in the total number of firms per worker could be 5 percentage points lower in 2012 or only about 15% of the total growth.

Two other details of the methodology used to produce NES data are worth mentioning. First, inclusion in the NES data is also subject to maximum revenue threshold, which vary across sectors but also remain constant over time in nominal terms. This presumably biases downwards the increase in the number of nonemployers over time without affecting measures of the total number of firms, as nonemployers with revenue above these thresholds are categorized as employers and included in the BDS data. Further, censored firms are likely older, limiting the impact on our measures of how $\bar{z}$ evolves over time that depend on the size of incumbents relative to entrants.

Second, the main criteria for inclusion in the NES data is that the firm be subject to federal income taxes. Therefore, non-profit nonemployers are not included. Given the rise in the number of non-profit firms in the U.S. over time, this could bias downwards our measures of the growth in the number of nonemployers and hence total number of firms.

D Sectoral Composition in the 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.[14] We find that the most significant change is

[14]We do not have data identifying government workers by industry, hence sectoral measures of employment include these workers.

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 D.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

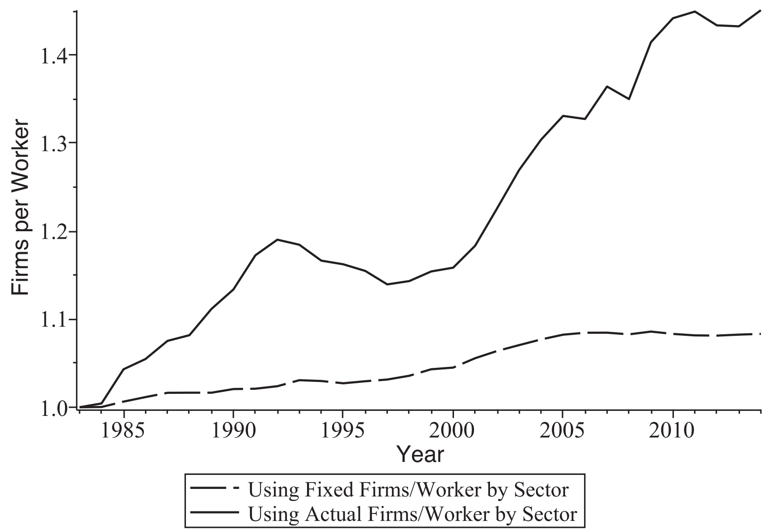
Notes: Data from Nonemployer Statistics (NES), Business Dynamics Statistics (BDS), Survey of Business Owners, and Statistical Abstract, U.S. Census Bureau, and Current Population Survey (CPS), U.S. Bureau of Labor Statistics.

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 D.2 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.

E Model Extension with Overhead Costs

In the model, we assume entry costs are the only non-production costs incurred by firms. In this appendix, we extend the model to show that allowing for realistically-specified overhead

Figure D.2: 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 (sectoral) 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. Data from Nonemployer Statistics (NES), Business Dynamics Statistics (BDS), Survey of Business Owners, and Statistical Abstract, U.S. Census Bureau, and Current Population Survey (CPS), U.S. Bureau of Labor Statistics.

costs (and allowing these costs to change over time) does not affect the main implications of the model. We start by noting that specifying overhead costs as fixed in nature (unrelated to the productivity of a firm) is inconsistent with the presence of very small firms in the data, a point made by Hsieh and Klenow (2014), among others. As a result of this insight, the literature specifies overhead costs as increasing in firm productivity. We therefore follow Asker et al. (2014) in assuming that a producer with productivity z must incur an output cost equal to $z \cdot c_p \cdot Y/L$ each period in order to operate. Instead of equation (9), operating profits for a producer with productivity z can now be expressed as;

$$\pi(z) = Az\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}(1-\alpha) - z \cdot c_p \cdot Y/L = \frac{w}{\alpha} \cdot \left(\frac{z}{\bar{z}}\right)\left[(1-\alpha)\left(\frac{L}{N}\right) - \bar{z}c_p\right], \tag{E.1}$$

where the right-hand-side expression takes into account equilibrium values of w and Y/L. Note that the terms in square-brackets imply that a firm's decision about whether to operate or not is independent of productivity. If a firm chooses to operate, all firms do (conditional on the number of firms). Free entry ensures that the discounted value of expected lifetime profits (net of operating costs) at entry is exactly equal to the cost of entry, resulting in the following characterization of the (simplified) free entry condition in equation (13):

$$c_{E,t} + c_{p,t} \cdot \bar{z}_{ent} = (1-\alpha)\left(\frac{L_t}{N_t}\right)\frac{\bar{z}_{ent}}{\bar{z}_t}$$

$$+ \left(c_{E,t+1} + c_{p,t+1} \cdot \bar{z}_{ent}\right) \frac{(1-\lambda_{t+1})(1+g_{t+1})}{1+R_{t+1}} \left(\frac{Y_{t+1}/L_{t+1}}{Y_t/L_t}\right). \tag{E.2}$$

This characterization illustrates that including proportional overhead costs simply changes the interpretation of non-production costs from entry costs to the sum of entry and operating costs. Hence, with this specification, we conclude that non-production costs (rather than entry costs specifically) declined by 66% between 1982 and 2014. The model's implications for the impact of firm dynamics on aggregate TFP over time does not change.

F Implied Aggregate TFP with only Employers

We derive implications for the growth in aggregate productivity abstracting from nonemployers. We perform our analysis in Section 4.4 using only employer data from the BDS. BDS data report average employment for both entrant and incumbent employers. As a result, N^{emp}/L^{emp} and $\bar{z}^{emp}/\bar{z}^{emp}_{ent} = \ell^{emp}/\ell^{emp}_{ent}$ can be measured directly from the data for each year. Figure F.3 illustrates the implications for aggregate TFP.

Figure F.3: Implied Aggregate TFP with only Employers

(a) Decomposition: Average Productivity and Employer Firms/Worker

(b) Baseline with and without Nonemployers

Notes: Panel a: The solid line reports aggregate TFP over time implied by changes in $\bar{z}$, while the dashed line reports TFP over time implied by employers per worker, both derived using only employer data from the BDS. Panel b: The solid line reports changes in TFP implied by firm dynamics in the benchmark, while the dashed line reports the same using only employer data from the BDS.

Figure F.3a shows a slight decline (2.6%) in aggregate TFP over time due to the drop in the number of employers per worker. This is in contrast to the 7.1% increase in aggregate TFP due to the increase in the total number of firms per work when including nonemployers. The decline in aggregate TFP is more than offset by the increase in $\bar{z}$, which implies an increase of 4.8% in TFP from 1982 to 2014. Interestingly, the TFP effect associated with the increase in $\bar{z}$ is larger when including nonemployers (about 8% during the sample period). Figure F.3b shows a net increase over time of 2.1% in aggregate TFP implied by firm dynamics, much lower than the 15.6% increase when accounting for all firms, including nonemployers. The divergence in implied TFP when excluding nonemployers is more pronounced after 1997.

G Share of Nonemployers

To assess how much of the increase in nonemployers (as a share of all firms) is due to the overall increase in the number of firms, we perform a counterfactual exercise using our framework where we assume a constant productivity distribution across firms, a constant $\bar{z}$. We quantify the counterfactual change in the share of nonemployers. We proceed in four steps. First, we assume the underlying productivity distribution can be described by a Pareto distribution with scale parameter ξ, and obtain a value for ξ by targeting the revenue of the 43rd percentile firm relative to average revenue across all firms in 1987, using data from the 1987 SBO.[15] We obtain a value of $\xi = 1.04$. Second, we infer a value for z_1 in 1987 from this distribution by targeting the share of nonemployers N^{non}_{1987}/N_{1987} using the following equation implied by the Pareto distribution;

$$z_{1,1987} = \left(1 - \frac{N^{non}_{1987}}{N_{1987}}\right)^{-1/\xi}.$$

Third, we then infer z_1 for every other year using equation (22);

$$z_{1,t} = z_{1,1987} \cdot \left(\frac{N_t/L_t}{N_{1987}/L_{1987}}\right).$$

Finally, we calculate the counterfactual share of nonemployers as the resulting share of firms with $z < z_1$ in each year, given the shape of the productivity distribution. This share is equal to;

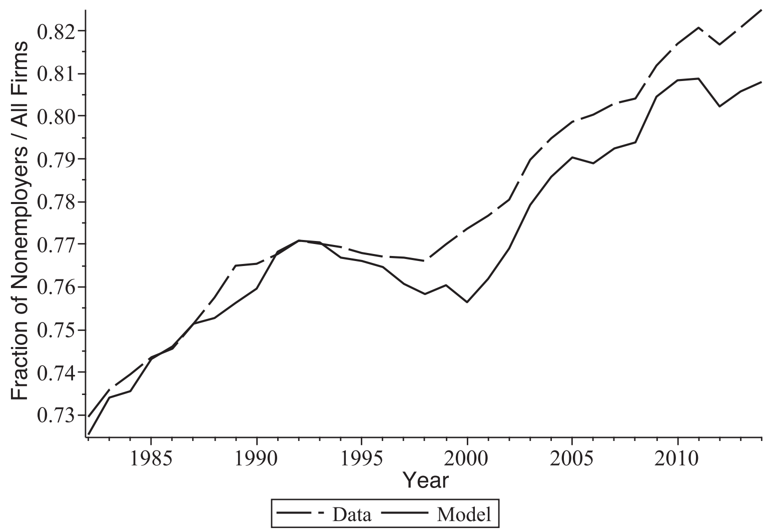
$$\frac{N^{non}_t}{N_t} = 1 - z_{1,t}^{-\xi}.$$

Figure G.4 compares our results to the observed share of nonemployers in the data. The counterfactual suggests that most of the observed increase in the share of nonemployers is accounted for by the increase in the number of firms per worker. While the data show an increase of 9.5 percentage points in the share (from 73% in 1982 to 82.5% to 2014), our counterfactual generates an increase of 8.3 p.p. increase.

We note that the above counterfactual experiment abstracts from potential interactions between

[15] 1987 is the earliest year this data is published for, and we choose the 43rd percentile because it is the closest reported percentile to the median.

Figure G.4: Nonemployer Share of All Firms



Notes: Data from Nonemployer Statistics (NES), Business Dynamics Statistics (BDS), Survey of Business Owners, and Statistical Abstract, U.S. Census Bureau. Model counterfactual infers how share would change if firm productivity distribution over time remained as in 1987.

changes in the number of firms per worker and changes in the relative productivity distribution across firms. We therefore perform another counterfactual where we fix the number of firms per worker, but allow the distribution of z to change. For this exercise, we narrow our focus between 1987 to 2012, the earliest and latest years for which we have the necessary data. With values for ξ and z_1 already in hand for 1987, we proceed in three steps. First, we similarly assume the productivity distribution across all firms in 2012 can also be described by a Pareto distribution with a time-specific shape parameter ξ_{2012}. We obtain a value for this parameter by targeting the revenue of the 55th percentile firm (the closest reported to the median) relative to average revenue across firms using data from the 2012 SBO. We obtain a value of $\xi_{2012} = 1.01$. Second, given our value for $z_{1,1987}$ above and assuming N/L remains constant at its 1987 level, we obtain a counterfactual value for $z_{1,2012}$ using equation (22);

$$z_{1,2012} = z_{1,1987} \cdot \frac{\bar{z}_{2012}}{\bar{z}_{1987}},$$

where values for $\bar{z}$ in 1987 and 2012 are taken from Section 4.3. Finally, we calculate the counterfactual share of nonemployers in 2012 as the resulting share of firms with $z < z_{1,2012}$, given ξ_{2012}. This share is equal to;

$$\frac{N_{2012}^{non}}{N_{2012}} = 1 - \left(z_{1,2012}\right)^{-\xi_{2012}}.$$

From 1987 to 2012, the observed nonemployer share of all firms in the data rises from 75.1% to 81.7%, a 6.6 p.p. increase. Figure G.4 shows the nonemployer share would have risen to 80.2% over the same period if the across-firm productivity distribution had remained constant, a 5.1 p.p. increase. If instead the number of firms per worker had remained fixed over time, our second counterfactual suggests the nonemployer share would have risen to 79.6%, a 4.5 p.p. increase. We conclude from these counterfactuals that both the change in the number of firms per worker and the change in the relative productivity distribution are important in accounting for the increase in the share of nonemployers we document in our sample period.