

# Firm Growth through New Establishments\*

Dan Cao<sup>†</sup>  
Georgetown University

Henry R. Hyatt<sup>§</sup>  
Census Bureau

Toshihiko Mukoyama<sup>‡</sup>  
Georgetown University

Erick Sager<sup>¶</sup>  
Federal Reserve Board

February 2019

## Abstract

This paper analyzes firm growth along two margins: the extensive margin (adding more establishments) and intensive margin (adding more workers per establishment). We utilize administrative datasets to document the behavior of these two margins in relation to changes in the U.S. firm size distribution. Between 1990 and 2015, we find that the significant increase in average firm size was driven primarily by the expansion along the extensive margin, particularly in superstar firms. We develop a general equilibrium model of endogenous innovation that features both extensive and intensive margins of firm growth. We estimate the model to uncover the fundamental forces that caused the changes over this time period through the lens of our model. We find that, over time, the cost of innovations that lead to new establishments has declined for firms who are innovative in that dimension. Meanwhile, the duration that a firm can enjoy high growth through such innovation became shorter, and firm entry became more costly.

*Keywords:* firm growth, firm size distribution, establishment, innovation

*JEL Classifications:* E24, J21, L11, O31

---

\*First version: February 2017. We thank Jess Helfand and Mike LoBue for their help with Quarterly Census of Employment and Wages data, helpful comments and suggestions from Stephane Bonhomme, John Earle, David Ratner and Mike Siemer. This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed within are those of the authors and do not necessarily reflect those of the Federal Reserve System, the Census Bureau, the Bureau of Labor Statistics or the U.S. government. All errors are our own.

<sup>†</sup>Contact information: dan.cao@georgetown.edu

<sup>‡</sup>Contact information: tm1309@georgetown.edu

<sup>§</sup>Contact information: henry.r.hyatt@gmail.com

<sup>¶</sup>Contact information: erick.sager@gmail.com

# 1 Introduction

Understanding the process of firm growth is essential in the analysis of macroeconomic performance. Firms that innovate and expand are the driving force of output and productivity growth. Recent analysis of macroeconomic productivity emphasizes the role of innovation and reallocation at the firm level, both from theoretical and empirical standpoints.

In this paper, we focus on a particular aspect of firm growth: growth through adding new establishments. In general, a firm can increase its size, measured in employment, along two margins. It can add more workers for given establishments, or build new establishments. We call the former the *intensive margin* and the latter the *extensive margin*. This distinction is important because these margins typically imply different reasons for expansion. In the context of manufacturing, a new plant often is built to produce a new product. In the service sector, building a new store or a new restaurant imply venturing into a new market. In both contexts, creating a new establishment requires a significant investment.

Given the importance of this distinction, it is surprising that very little is known about how firms grow through building establishments. With the exception of a few papers, the literature separately treats firms and establishments without linking the two. Our first goal is to establish stylized facts about how firms and establishments are linked. We then interpret the facts through the lens of a macroeconomic model of endogenous innovation and firm dynamics.

In the literature on firm dynamics, firms and establishments are often treated as interchangeable. One justification for this indifference is the fact that the majority of firms are single-establishment firms. While this assumption is justifiable in some situations, it is misleading in many macroeconomic contexts. In the U.S. economy, while 95% of firms are single-establishment firms, their share in total employment is less than half (45%).<sup>1</sup> Furthermore, the firm size distribution exhibits a Pareto tail, which implies that a large firm with many establishments has a disproportionately large impact on macroeconomic performance (Gabaix, 2011). Thus, understanding how these large firms are created is an important research question.

We find that the average firm size has grown in recent years, and it is largely driven by the firms in the right tail of the size distribution. This result echoes the recent emphasis on the emergence of superstar firms (Autor et al., 2017) and also consistent with the increase in concentration measured by the markup (De Loecker and Eeckhout, 2017). These studies suggest that this increase in size has had important implications on other changes in macroeconomic variables, such as the decline in labor share. What is novel about our empirical analysis is that we show that this expansion is driven by extensive margin growth.<sup>2</sup>

To investigate what changes in the economic environment have contributed to this phenomenon,

---

<sup>1</sup>The numbers are taken from 2005 Q1 Business Employment Dynamics (which is drawn from the Quarterly Census of Employment and Wages, explained in Section 2.1): <https://www.bls.gov/bdm/sizeclassqanda.htm>. In the manufacturing sector, single-plant firms own 72% of plants, they produce only 22% of the value added (Kehrig and Vincent, 2017).

<sup>2</sup>Choi and Spletzer (2012) and Hathaway and Litan (2014) also note the discrepancy in the recent trend between firm size and establishment size.

we build a macroeconomic model of endogenous firm growth. Our model extends previous work by Klette and Kortum (2004) and Luttmer (2011). In Klette and Kortum (2004) and Luttmer (2011), each individual firm grows by adding production units (*external innovations* in our terminology), “product lines” in Klette and Kortum’s (2004) terminology and “blueprints” in Luttmer’s (2011) terminology. A natural interpretation of such a production unit, as is explicit in Luttmer (2011), is an establishment. Thus, this type of framework provides a perfect vehicle for analyzing firm growth through new establishments. The major departure of our model, compared to Klette and Kortum (2004) and Luttmer (2011), is the recognition that each establishment can grow. In fact, the critique of Klette and Kortum (2004) by Acemoglu and Cao (2015) points out the lack of such growth that we observe in the data. We introduce this technological improvement at the establishment level (we call it *internal innovation*), and explicitly compare our model outcomes to the data on establishments.

We estimate model parameters to match the change in U.S. firm size distributions between 1990 and 2015, for both the intensive margin of workers per establishment and the extensive margin of the number of establishments within firms. The estimated model interprets the increase in average firm size and the significant increase in the number of establishments as caused mainly by (a) a reduction in costs of innovations that lead to creating new establishments for firms conducting such innovations, (b) a decline in the duration that a firm can enjoy high growth driven by a new establishment, and (c) an increase in entry costs. Taken together, the model shows that an increase in the size of large firms through establishment creation is accompanied by a decline in the firm entry rate, as has been noted by Decker, Haltiwanger, Jarmin, and Miranda (2014) and others.

The model constructed by Akcigit and Kerr (2016) has similar features to that in this paper. They also consider innovations that are internal to the establishments (“products” in their terminology) that the firm already produces and external innovations that increases the number of establishments that the firm operates. At the same time, many of their model assumptions are different from ours.<sup>3</sup> The differences in assumptions mainly stem from a difference in purpose for each model. Our main purpose is to map the model to the data on firm and establishment sizes, measured by employment. Akcigit and Kerr (2016) primarily use patent data and therefore consider “products” instead of “establishments.” Our model is tractable and it allows for analytical characterizations of Pareto tails in the firm size, the establishment size, and the number of establishments per firm. Our model is tailored to the question we ask: what is behind the increase in firm size over the recent years? In particular, we focus on the role of changes in innovation costs in allowing firms to expand on the extensive margin.

Lentz and Mortensen (2008) extends Klette and Kortum’s (2004) model and estimates it using Danish data. While their main focus is on productivity and reallocation, this paper mainly targets

---

<sup>3</sup>For example, in Akcigit and Kerr (2016) an external innovation improves a product that is already produced by another firm, while in our model it creates a new establishment whose quality is the same as the other establishments in the firm. Our assumption seems more appropriate especially in the context of service sector innovations. For example, when a retail firm opens a new store after researching the placement of a new location, the quality of the new store would more likely be associated with the firm itself than the existing stores in that location.

labor market facts. We directly exploit establishment-level information in analyzing firm dynamics. Furthermore, the empirical phenomenon we highlight mainly concerns large firms, and the model is tailored to fit the right tail of the firm size distribution.

The paper is organized as follows. Section 2 describes the empirical patterns of firm growth we find using our dataset. Section 3 sets up the model and Section 4 provides a theoretical characterization. Section 5 estimates model parameters and uses the model to perform a quantitative decomposition of 1990-2015 U.S. firm growth. Section 6 concludes.

## 2 Empirical facts

We first describe the data sources. Then we present cross-sectional and time series facts related to firm size, establishment size, and number of establishments per firm.

### 2.1 Data

Our data set is the Quarterly Census of Employment and Wages (QCEW). The microdata in the QCEW is collected for the official administration of state unemployment insurance programs, and therefore contains a near census of establishments in the United States. The QCEW reports payroll information on the number of workers and total wage bill by establishment and month between the years 1990-2016. Each establishment in the QCEW has an employer identification number (EIN) and a 6-digit NAICS code associated with its self-reported primary industry. This paper contains calculations from a sample of 38 states (including the District of Columbia, Puerto Rico and the U.S. Virgin Islands) that allow access to their confidential data through the Bureau of Labor Statistics' visiting researcher program.<sup>4</sup>

We use the employer identification numbers (EINs) as the definition of the firm. This is the level at which companies file their tax returns, and often considered as the boundary of the firm in recent studies. As is frequently discussed, EINs can be different from the ultimate ownership, especially for large firms. Song et al. (2016) discusses this at length in their study of inequality. They state that for 4,233 New York Stock Exchange publicly listed firms in the Dunn & Bradstreet database, 13,377 EINs are reported. For example, Walmart operates with separate EINs for "Walmart Stores," the Supercenter, Neighborhood Market, Sam's Club, and On-line divisions. Stanford University has separate EINs for the university, the bookstore, the main hospital, and children's hospitals. We view the EIN as a reasonable definition of a firm for our analysis, given that it is an economically meaningful unit (especially from the accounting perspective), and that typically the EINs for these large firms with multiple EINs are still substantially more aggregated than establishments.<sup>5</sup>

Further details of the dataset are described in Appendix A.

---

<sup>4</sup>Further details of the dataset are described in Appendix A, as well as previous work that uses the QCEW such as Ratner (2013), Siemer (2014), Chodorow-Reich (2014), and Chodorow-Reich and Wieland (2017).

<sup>5</sup>For example, as of October 31 2017, Walmart has 402 Discount Stores, 3552 Supercenters, 702 Neighborhood Markets, 660 Sam's Clubs, and 97 Small Formats including E-Commerce Acquisition/C-stores (numbers are taken from <https://corporate.walmart.com/our-story/locations/united-states>).

## 2.2 Cross-sectional characteristics

We start from the cross-sectional properties of firms and establishments. Although publicly-available datasets such as Business Dynamics Statistics of the U.S. Census Bureau contain size distribution of establishments and firms separately, it has not been documented how these are linked.

The particular interest of this paper is how firms grow in two margins, the intensive margin and the extensive margin. It is therefore the useful first step to describe the distributions of the firm size, the establishment size, and the number of establishment per firm. Figure 1 plots one minus cumulative distribution function in log-log scale, a type of figure commonly used in the literature to demonstrate whether the data exhibit Gibrat’s Law.<sup>6</sup>

To see how firm size is related to intensive and extensive margins, in Figure 2 we plot the average of establishment size at each firm (extensive margin) and the average of the number of establishment at each firm (intensive margin) in different firm size bins.<sup>7</sup> Denoting the firm size by  $Z$ , the number of establishment per firm as  $X$ , and the average establishment size for each firm as  $Y$ , one can decompose the firm size into extensive and intensive margin:

$$\log(Z) = \log(E[X|XY = Z]) + \log(E[Y|XY = Z]) + \Omega, \quad (1)$$

where  $\Omega \leq 0$  and it is equal to zero if and only if  $var[X|XY = Z] = 0$ .<sup>8</sup> In Figure 2, the left-hand side of (1) is the horizontal axis, and the first and the second terms are extensive and intensive margins. The sum of the slopes of the extensive and intensive margins can be different from one because of the  $\Omega$  term (which varies with  $Z$ ).<sup>9</sup>

## 2.3 Firm life cycle

It is well known that the firm growth exhibits a strong life cycle pattern. In Figure 3 we document firm growth patterns in terms of extensive and intensive margins, and additionally show these patterns for young and old firms.<sup>10</sup>

## 2.4 Changes of cross-sectional characteristics: 1990–2014

There have been notable changes in firm characteristics over our sample period. Figure 4 plots average firm size, measured by the number of workers within a firm. Over our sample period, the

---

<sup>6</sup>Due to a lengthy disclosure review process, we cannot include Figure 1 in the current version of this paper. An updated draft with a complete set of figures and tables is forthcoming pending review.

<sup>7</sup>Due to a lengthy disclosure review process, we cannot include Figure 2 in the current version of this paper. An updated draft with a complete set of figures and tables is forthcoming pending review.

<sup>8</sup>The derivation of (1) is in Appendix B.

<sup>9</sup>Xi (2016) draws a similar figure to Figure 2, although his graph describes different pattern for intensive margin averages. He is not explicit about how his graph is drawn (including the data source), but we obtain a similar graph to his once we put the second and the third term in (1) together and call it the intensive margin.

<sup>10</sup>Due to a lengthy disclosure review process, we cannot include Figure 3 in the current version of this paper. An updated draft with a complete set of figures and tables is forthcoming pending review.

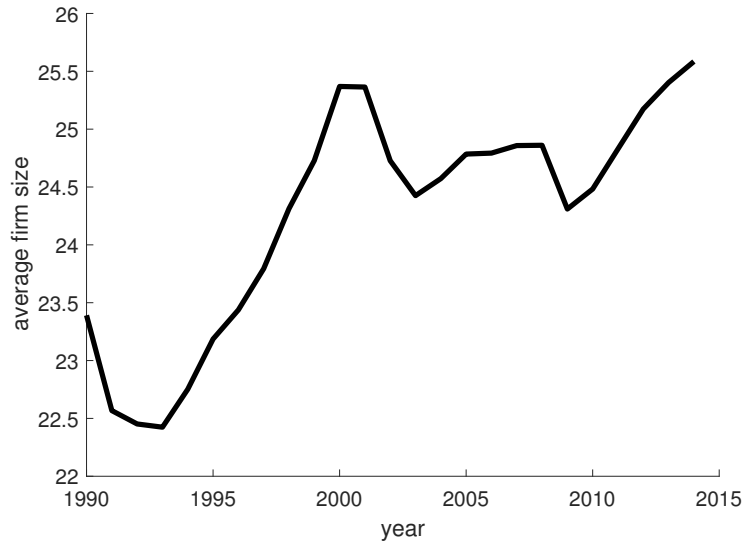


Figure 4: Average firm size (number of workers)

average firm size has increased from about 23 employees to over 25 employees. This fact accords well with the rise in concentration in the U.S. economy documented by, for example, Autor et al. (2017).

Figures 5 and 6 present the novel fact that we focus on in this paper.<sup>11</sup> Figure 5 plots the average establishment size of each firm, measured by employment.<sup>12</sup> It shows that, despite the increase in the firm size over our sample period, the establishment size remains stationary. In contrast, the average number of establishments per firm, shown in Figure 6, exhibits a strong upward trend. The average number of establishments per firm starts from 1.2 in 1990 and increases to over 1.5 in 2014. This contrasting behavior implies that different forces are at work for these different components of firm growth.

What drives the increase in firm size, in particular along the extensive margin? To investigate this phenomenon further in the micro level, we consider three dimensions of disaggregation: different sectors, different ages, and different size bins.

First, to see how the patterns are different across different sectors, Figure 7 plots the time series of the average firm size in each sector, compared to the size in 1990. It can be seen that all sectors experienced the size increase over the period of 1991-2013. Over the whole sample period, the service sector, which employs the majority of the U.S. workers, have experienced the largest size

<sup>11</sup>Choi and Spletzer (2012) and Hathaway and Litan (2014) also document trends in firm size and establishment size, but do not analyze the number of establishments per firm, which is central to our analysis.

<sup>12</sup>The average establishment size of each firm, calculated here, is a different concept from the average size of establishments in the whole economy. Here we first calculate the average establishment size within each firm. After that, we average this number over all firms. This is different from simply dividing the number of all workers by the number of the establishments in the whole economy. Therefore, in calculating the numbers here, having an access to the microdata is essential— this cannot be calculated from the publicly-available information of the total numbers of employment, the total number of firms, and the total number of establishments.

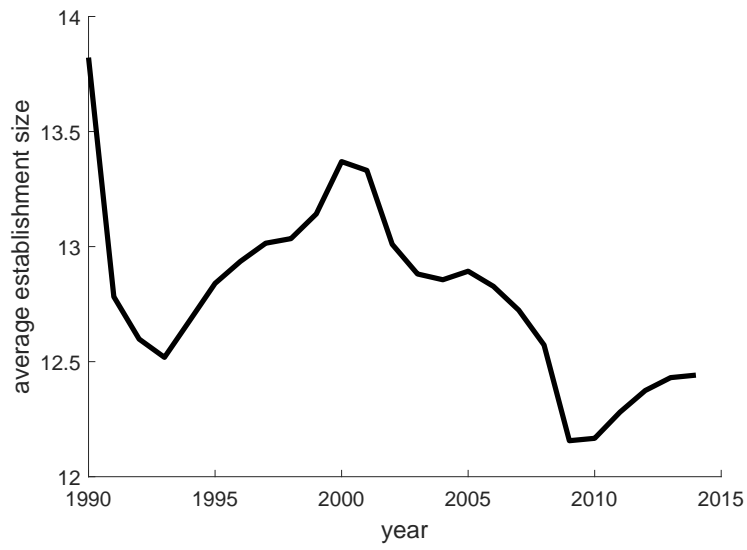


Figure 5: Average establishment size (number of workers)

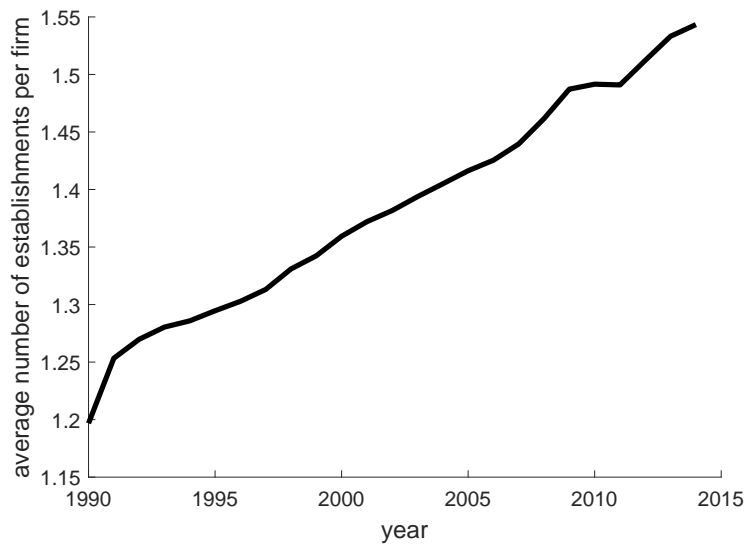


Figure 6: Average number of establishment per firm

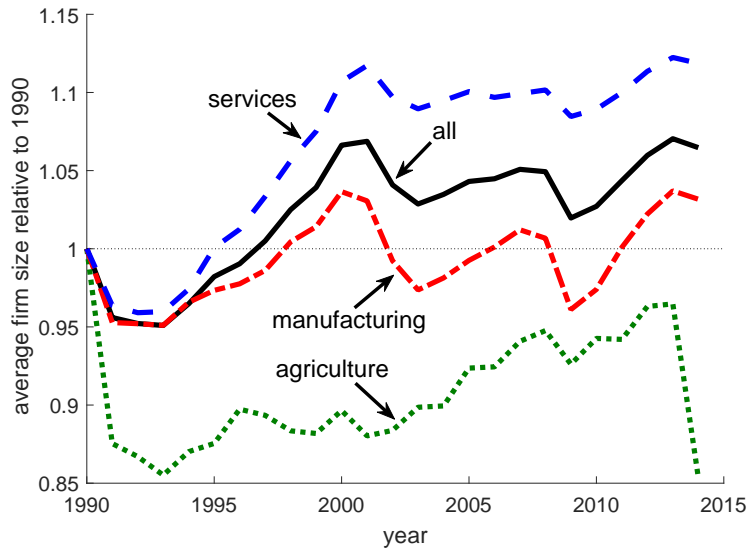


Figure 7: Average firm size (number of workers), different sectors

increase.

Similarly to the overall economy, the intensive margins plotted in Figure 8 exhibit a flat profile or a slight downward trend for all sectors. In contrast, Figure 9, which plot the extensive margin for different sectors, outlines the same message as the entire firm size: a significant increase in the expansion of service sector firms is the driving force of the overall firm size increase.

For the second dimension, we look at the outcomes for different ages of the firm. Note that the graphs for the age groups starts at the year 2001, so that we are able to consistently measure the age groups across different time periods. Figures 10, 11, and 12 repeat the same plots as earlier. Because the significant increase in the firm size occurs during 1990s, Figure 10 does not exhibit any obvious trend over the sample period for overall firms. However, we can see that there still is an increasing trend for the firms that are older than 11 years old. As in the case for all firms, the intensive margin does not exhibit an increasing trend in Figure 11 for any age group. In contrast, Figure 12 reveals a striking contrast across different age groups: the increase in extensive margin for overall economy is driven by the older firms. This motivates our modeling choice—firms are not born with different sizes across different time period; rather, their pattern of growth has changed over time.

To see another dimension of heterogeneity, Figure 13 calculates the average size within size bins. There is a pattern of spreading out: very small firms with 1 to 4 employees have tended to become smaller, while the average size of larger firms with 100 employees has increased over time. If we examine the very right tail of firms with 5000 workers or more, firm size has been increasing over time since 1997, with a similar increase to firms that have 100 employees or more.

Figure 14 looks at this for the intensive margin, and here we do not see an obvious pattern. None of the series have an increasing trend, and in fact, the overall time-series pattern looks similar



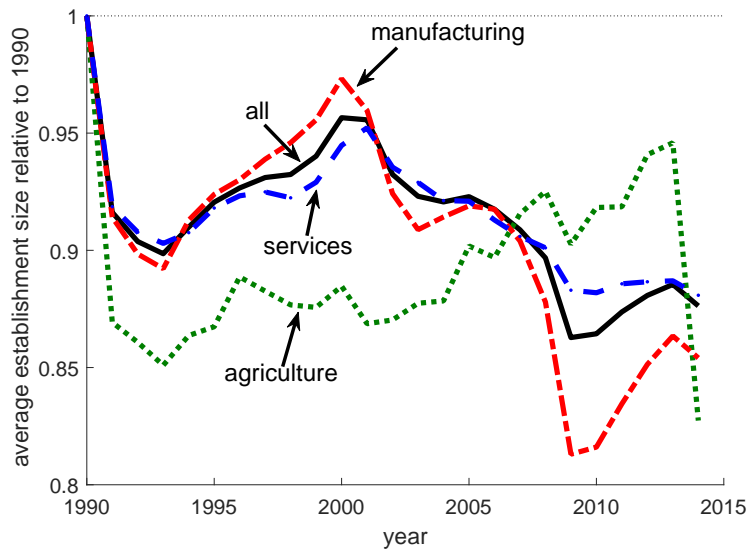


Figure 8: Average establishment size (number of workers), different sectors

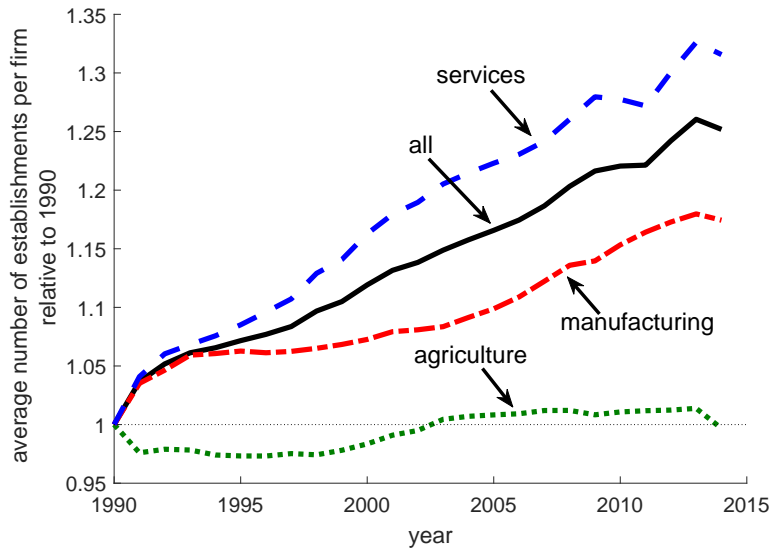


Figure 9: Average number of establishments per firm, different sectors

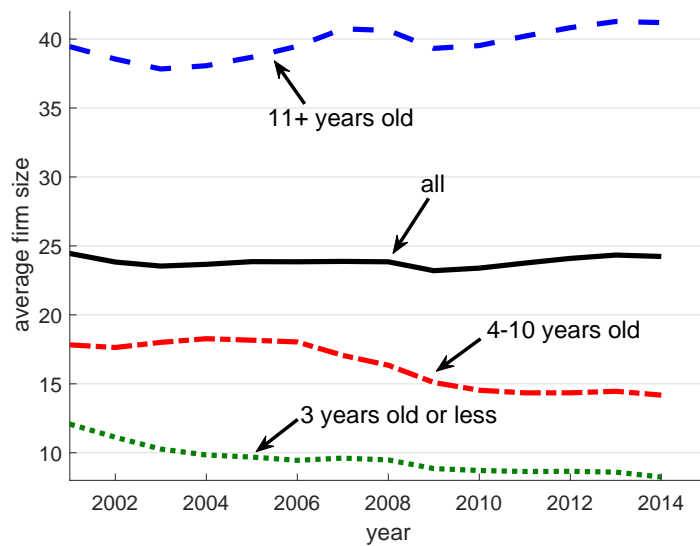


Figure 10: Average firm size (number of workers), different age groups

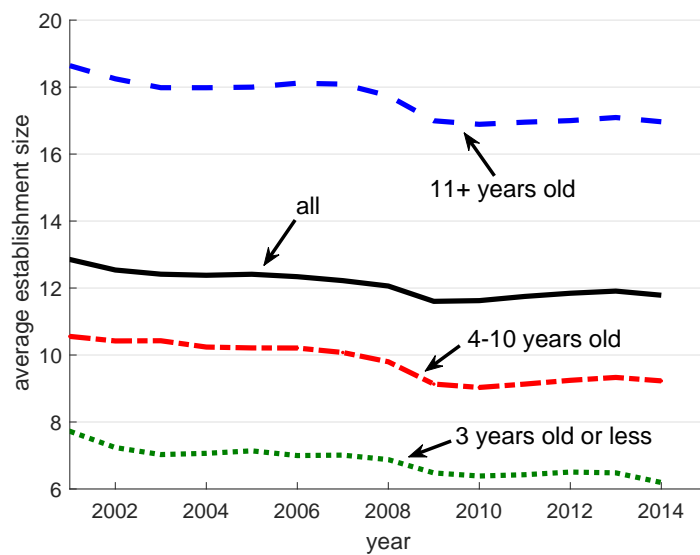


Figure 11: Average establishment size (number of workers), different age groups

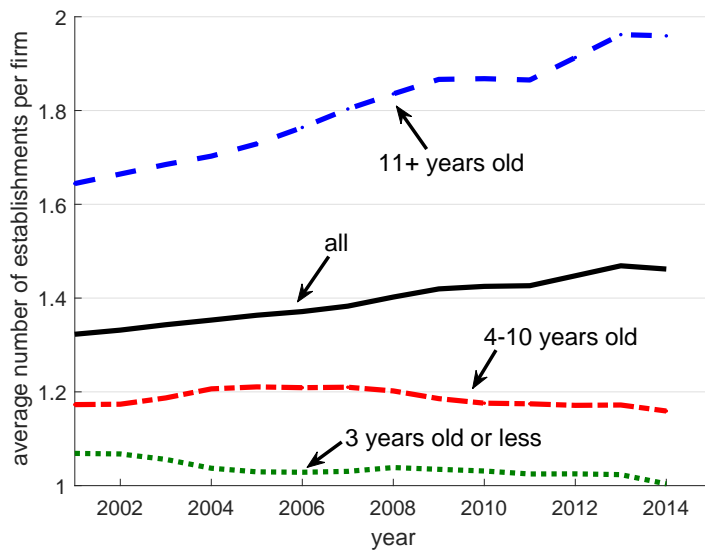


Figure 12: Average number of establishments per firm, different age groups

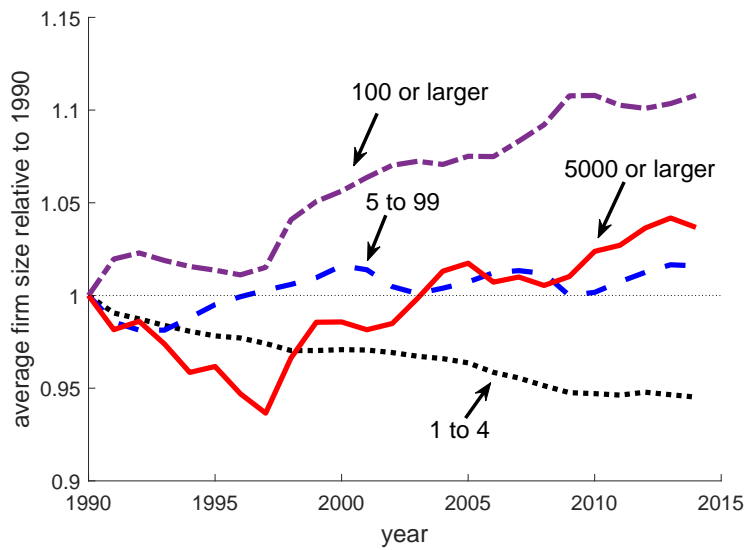


Figure 13: Average firm size (number of workers), different size bins

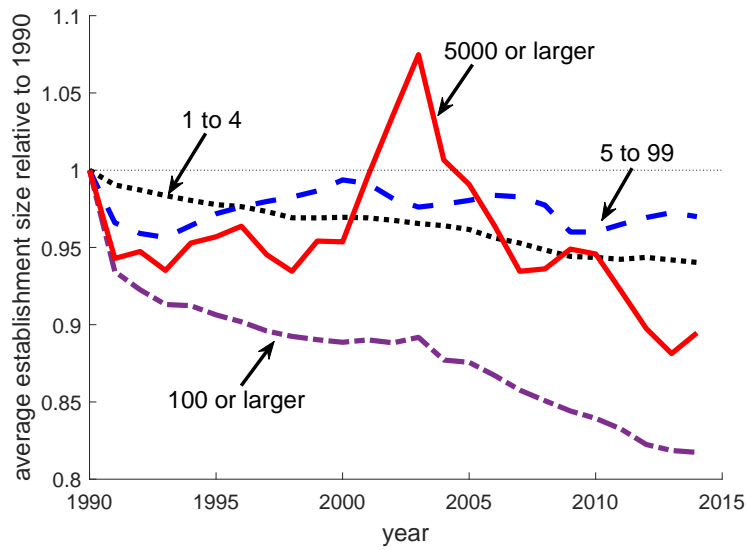


Figure 14: Average establishment size (number of workers), different size bins

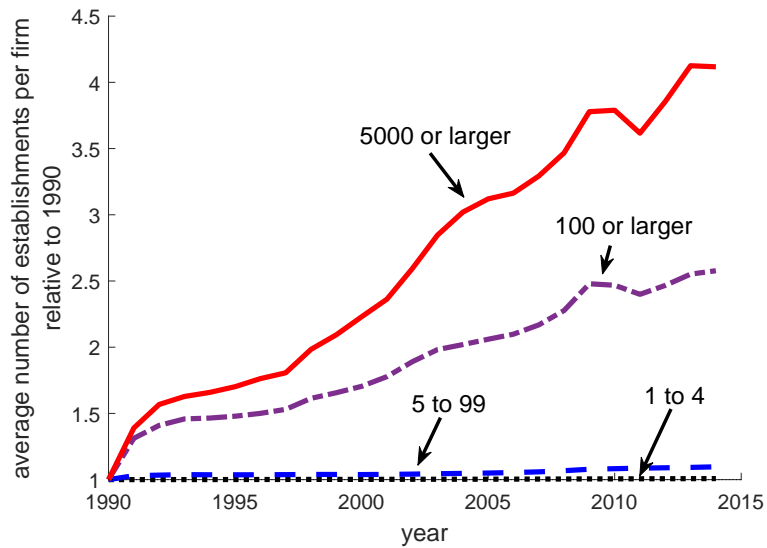


Figure 15: Average number of establishments per firm, different size bins

between very small firms (1 to 4 employees) and very large firms (5000 or more employees) except for a spike for very large firms in early 2000s.

The average number of establishment per firm, shown in Figure 15, reveals different trends for different firm sizes. Very small firms are predominantly single-establishment firms over the entire sample period. Medium-size firms with 5 to 99 employees have had a modest increase in the number of establishments. Larger firms have had a startling increase in the number of establishments. On average, the firms in 5000 or more employees category have about 4 times more establishments in 2014 compared to 1990. Thus, we conclude that a key mechanism that generated the increase in firm size in recent years is expansion through the number of establishments in very large firms.

### 3 Model

In this section, we construct a model of firm dynamics which can be mapped to our data observations. The heart of the model is the *endogenous* productivity improvements by intermediate-good firms, which allows us to analyze the fundamental causes of the recent changes in firm characteristics, which we described above.

#### 3.1 Model setting

Time is continuous. The representative consumer provides labor and consumes the final good. The final good is produced from differentiated intermediate goods.

#### 3.2 Representative consumer

The consumer side is intentionally kept simple, as we focus mainly on firm growth. The utility function of the representative households is

$$U = \int_0^\infty e^{-\rho t} L(t) \frac{(C(t)/L(t))^{1-\sigma}}{1-\sigma} dt.$$

The consumer consumes, owns firms, and supplies labor. The labor supply is given exogenously and grows at the rate  $\gamma \geq 0$ . Letting the real interest rate  $r$ , the Euler equation for the consumer is

$$\frac{\dot{C}(t)}{C(t)} = \frac{r - \rho}{\sigma}. \quad (2)$$

Final output is used for consumption, firm investments in innovative activities, and firm fixed costs:

$$Y(t) = C(t) + R(t) + K(t),$$

where  $Y(t)$  is the final goods output,  $R(t)$  is total R&D of the incumbents, and  $K(t)$  is the total entry cost.

### 3.3 Final good producers

There is a perfectly competitive final good sector. The final good is produced from differentiated intermediate goods. Intermediate goods have different qualities, and a high-quality intermediate good contributes more to the final good production. The production function for the final good is

$$Y(t) = \left( \int_{\mathcal{N}(t)} q_j(t)^\beta x_j(t)^{1-\beta} dj \right)^{\frac{1}{1-\beta}}, \quad (3)$$

where  $x_j(t)$  is the quantity of intermediate good  $j$ , and  $q_j(t)$  is its quality.  $\mathcal{N}(t)$  is the set of actively-produced intermediate goods and  $N(t) = |\mathcal{N}(t)|$  denote the number of actively-produced intermediate goods. We assume that  $\beta \in (0, 1)$ .

With the maximization problem

$$\max_{x_j(t)} \left( \int_{\mathcal{N}(t)} q_j(t)^\beta x_j(t)^{1-\beta} dj \right)^{\frac{1}{1-\beta}} - \int_{\mathcal{N}(t)} p_j(t) x_j(t) dj,$$

the inverse demand function for the intermediate good  $j$  is

$$p_j(t) = Y(t)^\beta \left( \frac{q_j(t)}{x_j(t)} \right)^\beta. \quad (4)$$

### 3.4 Intermediate good producers

The intermediate good sector is monopolistically competitive. Each intermediate good is produced by one firm. A firm can potentially produce many intermediate goods. A firm can add a new intermediate good to its portfolio by investing in R&D (external innovation). It can also increase the quality of the intermediate goods that it already produces by investing in R&D (internal innovation). A new firm can enter by coming up with its first product by innovation. Later we map one intermediate good to one establishment in the data.

We assume that the intermediate goods are produced only by labor. This is the only process in the entire economy that uses labor as an input. This allows us to map the employment dynamics of the intermediate good sector to our data analysis in Section 2.<sup>13</sup> The production function for intermediate good  $j$  is

$$x_j(t) = Z(t) \ell_j(t), \quad (5)$$

where  $Z(t)$  is the labor productivity. Here, we allow a geometric weighted average of exogenous factor, at rate  $\theta$  and endogenous factor,  $Q(t)$ :

$$Z(t) = (e^{\theta t})^\alpha Q(t)^{1-\alpha}, \quad (6)$$

---

<sup>13</sup>A similar idea of mapping the employment process to the productivity process is employed by Hopenhayn and Rogerson (1993), Garcia-Macia et al. (2016), and Mukoyama and Osotimehin (2018).

where

$$Q(t) \equiv \frac{1}{N(t)} \int_{N(t)} q_j(t) dj$$

is the average quality of intermediate goods.

Given final good producer's demand for its output (4), the intermediate goods producer's profit maximization results in standard optimal pricing with a markup over marginal cost:

$$p_j(t) = \frac{1}{1 - \beta} \frac{w(t)}{Z(t)}.$$

The optimal price together with (4) and (5) implies that the labor demand is proportional to  $q_j(t)$  given the aggregate variables:

$$\ell_j(t) = (1 - \beta)^{\frac{1}{\beta}} \bar{w}(t)^{-\frac{1}{\beta}} \frac{q_j(t)}{Z(t)Y(t)^{\frac{\beta}{1-\beta}}} \quad (7)$$

where  $\bar{w}(t) \equiv w(t)/(Z(t)Y(t)^{\frac{\beta}{1-\beta}})$  is the normalized wage.

Because firms (and establishments) differ only in their level of quality  $q_j(t)$ , employment per establishment varies proportionally to  $q_j(t)$  in the cross section. Similarly, profit is also proportional to  $q_j(t)$ :

$$\pi_j(t) = \bar{\pi}(t)q_j(t),$$

where

$$\bar{\pi}(t) \equiv \beta(1 - \beta)^{\frac{1-\beta}{\beta}} \bar{w}(t)^{\frac{\beta-1}{\beta}}. \quad (8)$$

**Innovation:** Innovations are carried out through R&D activity. The input for R&D is in final goods. For an existing intermediate-good firm, there are two kinds of innovations: internal innovation and external innovation.

Internal innovation raises the quality of the goods that a firm already produces. The total intensity of internal innovation is denoted by  $Z_{I,j}(t)$ . The innovation intensity per good is  $z_{I,j}(t) \equiv Z_{I,j}(t)/n_j(t)$ , where  $n_j(t)$  is the number of goods firm  $j$  produces. Then the quality improves according to the law of motion

$$\frac{dq_j(t)}{dt} = z_{I,j}(t)q_j(t).$$

Here we index  $q$  only by  $j$  because, as we will explain momentarily, the quality of all goods produced by firm  $j$  is always the same within the firm.

We assume that different firms can have different costs for innovation. In particular, we partition firms into different (finite) types, and assume that different types have different costs for innovation. We will detail later how types evolve over time. We denote the number of types by  $T$  and index the types by  $\tau$ . The R&D cost for internal innovation is assumed to be  $\mathbf{R}_I^\tau(Z_{I,j}(t), n_j(t), q_j(t))$ . As in Klette and Kortum (2004), we assume that the R&D cost function  $\mathbf{R}_I^\tau(Z_{I,j}(t), n_j(t), q_j(t))$  exhibits constant returns to scale with respect to  $Z_{I,j}(t)$  and  $n_j(t)$ . Then the R&D cost per good

can be denoted as

$$R_I^\tau(z_{I,j}(t), q_j(t)) \equiv \mathbf{R}_I^\tau(z_{I,j}(t), 1, q_j(t)) = \frac{\mathbf{R}_I^\tau(Z_{I,j}(t), n_j(t), q_j(t))}{n_j(t)}.$$

We further assume that

$$R_I^\tau(z_{I,j}(t), q_j(t)) = h_I(z_{I,j}(t))q_j(t)$$

for a strictly convex function  $h_I^\tau(\cdot)$ .

External innovation adds brand-new intermediate goods to the production portfolio of the firm. We assume that the new good has the same quality as the average quality of the goods produced by that firm. Thus, all products that firm  $j$  produces always have the same quality. The total intensity of external innovation is denoted by  $Z_{X,j}(t)$ . The innovation intensity per good is  $z_{X,j}(t) \equiv Z_{X,j}(t)/n_j(t)$ . The R&D cost for external innovation is assumed to be  $\mathbf{R}_X^\tau(Z_{X,j}(t), n_j(t), q_j(t))$ , which is assumed to be constant returns to scale with respect to  $Z_{X,j}(t)$  and  $n_j(t)$ . Once again, we can denote the cost per good as

$$R_X^\tau(z_{X,j}(t), q_j(t)) \equiv \mathbf{R}_X^\tau(z_{X,j}(t), 1, q_j(t)) = \frac{\mathbf{R}_X^\tau(Z_{X,j}(t), n_j(t), q_j(t))}{n_j(t)},$$

and we assume that

$$R_X^\tau(z_{X,j}(t), q_j(t)) = h_X^\tau(z_{X,j}(t))q_j(t)$$

for a strictly convex function  $h_X^\tau(\cdot)$ .<sup>14</sup>

**Dynamic Program:** We assume that firms transition between different types from  $\tau$  to  $\tau'$  with Poisson transition rates  $\lambda_{\tau\tau'}$ . Each establishment depreciates (is forced to exit) with the Poisson rate  $\delta_\tau$ . We also impose an exogenous exit shock at the firm level. Let  $d_\tau$  be the Poisson rate of the firm exit shock for a type- $\tau$  firm.

As in Klette and Kortum (2004), the Hamilton-Jacobi-Bellman (HJB) equation for the firm can be written separately for each establishment. The establishment-level HJB equation of a type- $\tau$  firm is

$$rV_\tau(q) - \dot{V}_\tau(q) = \max_{z_I, z_X} \left[ \begin{aligned} &\pi(q) - R_I^\tau(z_I, q) - R_X^\tau(z_X, q) + z_I \frac{\partial V_\tau(q)}{\partial q} q \\ &+ z_X V_\tau(q) - (\delta_\tau + d_\tau)V_\tau(q) + \sum_{\tau'} \lambda_{\tau\tau'} (V_{\tau'}(q) - V_\tau(q)) \end{aligned} \right],$$

where  $V_\tau(q)$  is the value of type- $\tau$  establishment with quality  $q$  and  $\dot{V}_\tau(q)$  is the time derivative of  $V_\tau(q)$  function. We omit the time notation here, because all variables and functions here are constant over time along the balanced-growth path.

As in Mukoyama and Osotimehin (2016),  $V_\tau(q)$  can be shown to be linearly homogeneous in  $q$  along the balanced-growth path. That is,  $V_\tau(q) = v_\tau q$  for a constant  $v_\tau$ . The HJB equation above

---

<sup>14</sup>A major departure from Luttmer (2011) is that we allow for internal innovation. Internal innovation allows us to capture the characteristics of intensive margin growth; we have shown in Section 2 that the size of establishments grow over time as firms age.



can be normalized to

$$rv_\tau = \max_{z_I, z_X} \left[ \bar{\pi} - h_I^\tau(z_I) - h_X^\tau(z_X) + (z_I + z_X - \delta_\tau - d_\tau)v_\tau + \sum_{\tau'} \lambda_{\tau\tau'}(v_{\tau'} - v_\tau) \right]. \quad (9)$$

Here,  $\bar{\pi}$  is given by (8). We assume that  $r - z_I - z_X + \delta_\tau + d_\tau > 0$ .

The HJB equation (9) implies that the choice of innovation intensities  $(z_I, z_X)$  is a function of the firm type only. We denote the decision rules as  $(z_I^\tau, z_X^\tau)$ . Recall that, from (8),  $\bar{\pi}$  is a function of  $\bar{w}$  only. Thus, given  $r$  and  $\bar{w}$ , the equation (9) and the first-order conditions can solve for  $v_\tau$ ,  $z_I^\tau$ , and  $z_X^\tau$ .

### 3.5 Entry

An intermediate firm can enter by creating a new product. A new firm draws its type from an exogenous distribution. Let the probability that an entrant draws the type  $\tau$  be  $m_\tau$ . Given the type  $\tau$ , the entrant draws the relative quality  $\hat{q}$ , which is equal to  $q(t)/Q(t)$ , from a distribution  $\Phi_\tau(\hat{q})$ . The value for entry  $V^e(t)$  is thus

$$V^e(t) = \sum_{\tau} m_\tau \int V_\tau(\hat{q}Q(t)) d\Phi_\tau(\hat{q}).$$

We assume that any potential entrant can pay a cost  $\phi Q(t)$ , denominated in final goods, to begin production. Therefore the free-entry condition is:  $V^e(t) = \phi Q(t)$ . By defining the value of entry relative to average product quality,  $v^e \equiv V^e(t)/Q(t)$ , we can rewrite the value of entry as

$$v^e = \sum_{\tau} v_\tau m_\tau \int \hat{q} d\Phi_\tau(\hat{q}). \quad (10)$$

and and free-entry condition as

$$v^e = \phi. \quad (11)$$

Note that once  $r$  is given, we can find a value of  $\bar{w}$  that satisfies the free entry condition (11). Let the number of entrants at time  $t$  be  $\mu_e N(t)$ , where  $\mu_e$  is a constant along the balanced-growth path.

### 3.6 Balanced growth path equilibrium

A *competitive equilibrium* of this economy is a wage  $w(t)$ , a consumer allocation  $(C(t), R(t), K(t))$ , a final good producer allocation  $(Y(t), \{x_j(t)\}_{j \in \mathcal{N}(t)})$ , an allocation for intermediate goods producers  $(\{\ell_j(t), q_j(t), p_j(t), z_{I,j}(t), z_{X,j}(t), n_j(t)\}_{j \in \mathcal{N}(t)})$  and a value of entry  $V_e(t)$  such that at each instant (i) consumers optimize, (ii) the final good producer's allocation solves its profit maximization problem, (iii) the intermediate good producers allocations solve their profit maximization problem, (iv) the free entry condition holds (v) the final good market clears,  $Y(t) = C(t) + R(t) + K(t)$ , and, (vi) the labor market clears  $L(t) = \int_{\mathcal{N}_t} \ell_j(t) dj$ .

We now construct a *balanced growth path* of this economy. Assume that the population  $L(t)$  grows at an exogenous rate  $\gamma$ . Furthermore, let aggregate quality  $Q(t)$  grow at a constant rate  $\zeta$ , and the number of establishments  $N(t)$  grow at a constant rate  $\eta$ . Denote the growth rate of final output  $Y(t)$  by  $g$ . Along a balanced growth path, the growth rates of  $Y(t)$ ,  $C(t)$  and  $R(t) + K(t)$  must all be equal.

From the Euler equation (2), we know that  $\dot{C}(t)/C(t) = g$  because  $C(t)$  grows at the same rate as  $Y(t)$ . This implies that  $r = \rho - \sigma g$  along the balanced growth path, so that we can calculate  $r$  once we know  $g$ .

Along the balanced-growth path, the quality-invariant component of profit  $\bar{\pi}(t)$  in (8) is constant. Therefore,  $w(t)$  must grow at the same rate as  $Z(t)Y(t)^{\frac{\beta}{1-\beta}}$ . Given that  $Y(t)$  grows at rate  $g$  and  $Z(t)$  grows at rate  $\alpha\theta + (1 - \alpha)\zeta$  from (6), then  $w(t)$  must grow at the rate  $\beta g/(1 - \beta) + \alpha\theta + (1 - \alpha)\zeta$ . This implies that the labor income of the representative consumer,  $w(t)L(t)$ , grows at the rate of wages plus population growth. Since labor income must grow at the same rate as consumption, we know that

$$g = \gamma + \alpha\theta + (1 - \alpha)\zeta + \frac{\beta}{1 - \beta}g \quad (12)$$

must hold.

Next, we can use the labor market clearing condition to further refine the expression for the growth rate for final output. Total labor demand can be calculated from (7),

$$\int_{\mathcal{N}(t)} \ell_j(t) dj = \left( \frac{w(t)}{(1 - \beta)Z(t)} \right)^{-\frac{1}{\beta}} \frac{Y(t)N(t)Q(t)}{Z(t)}. \quad (13)$$

The labor market clearing condition  $L(t) = \int_{\mathcal{N}_t} \ell_j(t) dj$  then yields the following expression for  $\bar{w}(t)$ ,

$$\bar{w}(t)L(t) = (1 - \beta) \left( \frac{N(t)Q(t)}{Z(t)Y(t)^{\frac{1}{1-\beta}}} \right)^{\beta} L(t)^{1-\beta}. \quad (14)$$

Since we just showed that the normalized wage does not grow along the balanced growth path, the above expression (14) implies

$$0 = \beta \left( -\gamma + (\eta + \zeta) - (\alpha\theta + (1 - \alpha)\zeta) - \frac{\beta}{1 - \beta}g \right),$$

which can be combined with condition (12) to yield

$$g = \eta + \zeta. \quad (15)$$

Hence, the growth of aggregate output is driven by the two margins of firm growth: the growth of the number of establishments  $N(t)$  and the growth of the average quality of products  $Q(t)$ .

Furthermore, we can rewrite (12) to obtain an explicit formula for  $\zeta$  given  $g$ :

$$\zeta = \frac{1}{1-\alpha} \left( \frac{1-2\beta}{1-\beta} g - \gamma - \alpha\theta \right).$$

Because  $\eta = g - \zeta$ ,

$$\eta = g - \frac{1}{1-\alpha} \left( \frac{1-2\beta}{1-\beta} g - \gamma - \alpha\theta \right) \quad (16)$$

holds.

## 4 Characterization of the model

In order to inform our understanding of the tradeoffs underlying the mechanisms of this model, we analytically characterize firm decisions, aggregate growth and the distributions of establishment number and size. This characterization will be useful for anchoring the interpretation of quantitative experiments in Section 5.

### 4.1 Output growth

We first derive the growth rate of final output from disaggregated firm behavior as follows. First note that the total number of establishments at each type (denote it as  $N_\tau(t)$  for type  $\tau$ ) has to grow at the same rate,  $\eta$ . In other words,  $N_\tau(t) = M_\tau N(t)$  has to hold for a constant  $M_\tau$  that satisfies  $M_\tau \in [0, 1]$  for all  $\tau$  and

$$\sum_{\tau} M_\tau = 1. \quad (17)$$

The law of motion for  $N_\tau(t)$  is

$$\dot{N}_\tau(t) = z_X^\tau N_\tau(t) - (\delta_\tau + d_\tau) N_\tau(t) + \mu_e m_\tau N(t) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} N_\tau(t) + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} N_{\tau'}(t).$$

The first term is the increase in products due to external innovation, the second term is the exit, the third term is the entry, and the fourth and fifth terms are the changes in firm types. The growth rate of  $N_\tau(t)$  has to be the common value  $\eta$ . Thus this equation can be rewritten as

$$\eta = z_X^\tau - (\delta_\tau + d_\tau) + \mu_e \frac{m_\tau}{M_\tau} - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{M_{\tau'}}{M_\tau}. \quad (18)$$

A simpler expression for  $\eta$  can be found by multiplying  $M_\tau$  to both sides of (18) and adding up:

$$\eta = \sum_{\tau} M_\tau [z_X^\tau - (\delta_\tau + d_\tau)] + \mu_e.$$

Note that using (16), (18) can be rewritten as

$$g - \frac{1}{1-\alpha} \left( \frac{1-2\beta}{1-\beta} g - \gamma - \alpha\theta \right) = z_X^\tau - (\delta_\tau + d_\tau) + \mu_e \frac{m_\tau}{M_\tau} - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{M_{\tau'}}{M_\tau}. \quad (19)$$

For notational convenience, define

$$Q_\tau(t) \equiv \frac{1}{N_\tau(t)} \int_{\mathcal{N}_\tau(t)} q_j(t) dj,$$

where  $\mathcal{N}_\tau(t)$  is the set of actively produced goods by type- $\tau$  firms. Balanced growth implies that  $Q_\tau(t)$  has to grow at the same rate as  $Q(t)$ . Also define

$$s_\tau \equiv M_\tau Q_\tau(t) / Q(t). \quad (20)$$

Note that  $s_\tau$  is constant along the balanced-growth path and it satisfies

$$\sum_\tau s_\tau = 1. \quad (21)$$

Using these notations, we can show that for any  $\tau$ ,  $g$  satisfies

$$g = z_I^\tau + z_X^\tau - (\delta_\tau + d_\tau) + \mu_e \frac{m_\tau}{s_\tau} \int \hat{q} d\Phi(\hat{q}) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{s_{\tau'}}{s_\tau}. \quad (22)$$

The derivation is in Appendix B. A simpler expression for  $g$  can be found by multiplying  $s_\tau$  on both sides of (22) and summing across  $\tau$ :

$$g = \sum_\tau s_\tau [z_I^\tau + z_X^\tau - (\delta_\tau + d_\tau)] + \mu_e \int \hat{q} d\Phi(\hat{q}).$$

The first term is the incumbent firms' contribution to  $g$  and the second term comes from entrants.

## 4.2 Properties of the model with one type

Although the quantitative exercise below features multiple types, the case with one type allows us sharper characterizations of the general equilibrium and help us develop intuitions. We omit the subscripts and superscripts  $\tau$  as there is only one type.

Recall the output growth rate  $g$  can be decomposed into the growth of  $Q(t)$  and the growth of  $N(t)$  with the formula (15). The separate formula for  $\zeta$  and  $\eta$  are particularly simple for the single-type case:

$$\zeta = z_I + \mu_e \left( \int \hat{q} d\Phi(\hat{q}) - 1 \right) \quad (23)$$

and

$$\eta = z_X - \delta - d + \mu_e \quad (24)$$

The sum,  $g$ , is

$$g = z_I + z_X - \delta - d + \mu_e \int \hat{q} d\Phi(\hat{q}). \quad (25)$$

The single-type case is particularly convenient as it has the recursive structure in solution. The free-entry condition (11) and (10) imply that

$$v = v_e = \phi.$$

From the first-order condition for innovations, this imply that  $z_I$  and  $z_X$  are determined only by  $\phi$  and the innovation cost functions  $h_I(\cdot)$  and  $h_X(\cdot)$ . Then the only endogenous variable to determine  $\zeta$  and  $g$  in (23) and (25) is  $\mu_e$ . Plugging (23) and (25) into (12) yields an equation with one unknown  $\mu_e$ . With some algebra, this equation becomes

$$\left[ \left( \frac{1-2\beta}{1-\beta} - (1-\alpha) \right) \int \hat{q} d\Phi(\hat{q}) + 1 - \alpha \right] \mu_e + \left[ \frac{1-2\beta}{1-\beta} - (1-\alpha) \right] z_I + \frac{1-2\beta}{1-\beta} (z_X - \delta - d) = \gamma + \alpha\theta.$$

We assume that  $\beta$  and  $\alpha$  are sufficiently small so that the coefficients of  $\mu_e$ ,  $z_I$ , and  $(z_X - \delta - d)$  on the left-hand side are all positive. Then the above steps solve out  $\mu_e$ ,  $\zeta$ ,  $g$ ,  $\eta$ ,  $z_I$ ,  $z_X$ , and  $v$  as functions of parameters. The only equilibrium object left here is  $\bar{w}$ , which can be solved by the HJB equation

$$(\rho + \sigma g)v = \bar{\pi} - h_I(z_I) - h_X(z_X) + (z_I + z_X - \delta - d)v$$

and the relationship (8).

The following propositions are straightforward and therefore presented without proof.

**Proposition 1** *An increase in the entry cost,  $\phi$ , increases  $z_I$  and  $z_X$ , while reducing  $\mu_e$ . When  $\alpha = 1$  (exogenous growth), these effects exactly offset and  $g$  stays constant. When  $\alpha < 1$  and  $\int \hat{q} d\Phi(\hat{q}) \leq 1$ ,  $g$  increases.*

**Proposition 2** *Suppose that the innovation cost functions take the form*

$$h_i(z_i) = \chi_i z_i^\psi,$$

*where  $i = I, X$ . The parameters satisfy  $\chi_i > 0$  and  $\psi > 1$ . Then a decrease in  $\chi_I$  increases  $z_I$  but keeps  $z_X$  the same. The entry rate  $\mu_e$  decreases, and when  $\alpha < 1$ , the overall growth rate increases. A decrease in  $\chi_X$  increases  $z_X$  but keeps  $z_I$  the same. The entry rate  $\mu_e$  decreases, and when  $\alpha < 1$  and  $\int \hat{q} d\Phi(\hat{q}) \leq 1$ ,  $g$  increases.*

In above Propositions, the direct effect through the changes in  $z_I$  and  $z_X$  have stronger impact on growth than the (offsetting) effect of  $\mu_e$ , unless  $\int \hat{q} d\Phi(\hat{q}) > 1$ .

**Proposition 3** *An increase in  $\delta$ ,  $d$ ,  $\gamma$ , or  $\theta$  keeps  $z_I$  and  $z_X$  constant, while increasing  $g$  through an increase in  $\mu_e$ . Changes in  $\alpha$  and  $\beta$  do not affect  $z_I$  and  $z_X$  either, although they influence  $g$  through the change in  $\mu_e$ .*

**Proposition 4** *Changes in the preference parameters  $\sigma$  and  $\rho$  do not have any effects on  $z_I$ ,  $z_X$ , and  $g$ .*

From above analysis, we can suspect that the increase in the number of establishments per firm, which is likely be driven by an increase in  $z_X$ , can be explained by the decrease in  $\chi_X$ . Other changes in parameters either move  $z_I$  and  $z_X$  together or move only  $\mu_e$  without affecting  $z_X$ . The slowdown in overall growth cannot be accounted for by the change  $\chi_X$ . Additional change in other technology parameters, such as  $\delta$ ,  $d$ ,  $\gamma$ , or  $\theta$ , can account for the overall growth slowdown.

### 4.3 Distributions of firm sizes and establishment sizes

The properties of our model allow us to analyze the firm size distribution from two different margins. In one margin, the number of establishments per firm evolves through the external innovation and exit shock. In the other margin, the size of each establishment evolves through the internal innovation. Note that the existence of these two margins is the major departure from Klette and Kortum (2004) and Luttmer (2011). In these papers, establishments are homogeneous and each establishment does not grow, so that the only relevant innovation is the external one.

Before analyzing the details, first note one general property of the model. The model assumptions imply that the establishments are homogeneous within a firm. A firm starts with one establishment, and whenever it expands the number of establishments, a new establishment inherits the same quality as the existing establishments. The intensity of the internal innovation,  $z_I^T$ , is common across establishments within a firm, while it may change over time. These two facts mean that establishments share common quality within a firm at any point in time. This also imply that the establishment sizes are also common within a firm. Although in reality establishment sizes are not the same within a firm, we view this as a useful simplification. In a balanced-growth equilibrium, the number of firms grow at the same rate as  $N(t)$ . The distribution of the number of establishments per firm,  $n(t)$ , is stationary, and the average quality of each establishment,  $q(t)$ , grows at the same rate as  $Q(t)$  from the definition of  $Q(t)$ . In the following, we provide further characterizations of the size distributions of establishments and firms. We assume that the economy is on a stationary balanced-growth path (BGP).

#### 4.3.1 General characterization

When there are multiple types of firms, it is substantially more complex to characterize the joint distribution of  $(n, \hat{q})$  for each types than the one-type case. Our approach here is to look at two separate margins: the distribution of the number of establishments per firm, where we denote the  $N(t)$ -normalized measure of type- $\tau$  firms with  $n(t) = n$  as  $\bar{\mathcal{M}}_\tau(n)$ , and the distribution of establishment quality, where we denote the fraction of type- $\tau$  establishments with  $q(t) \geq \hat{q}Q(t)$  as  $\bar{\mathcal{H}}_\tau(\hat{q})$ .

The first margin,  $\bar{\mathcal{M}}_\tau(n)$ , can be characterized by the following equations: for each  $\tau \in \Gamma$ :

$$\begin{aligned}
0 = & -(z_X^\tau + \delta_\tau + d_\tau + \eta)\bar{\mathcal{M}}_\tau(1) + 2\delta_\tau\bar{\mathcal{M}}_\tau(2) + \mu_e m_\tau \\
& - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \bar{\mathcal{M}}_\tau(1) + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \bar{\mathcal{M}}_{\tau'}(1)
\end{aligned} \tag{26a}$$

and for  $n > 1$ :

$$\begin{aligned}
0 = & -(n(z_X^\tau + \delta_\tau) + d_\tau + \eta)\bar{\mathcal{M}}_\tau(n) + (n+1)\delta_\tau\bar{\mathcal{M}}_\tau(n+1) + (n-1)z_X^\tau\bar{\mathcal{M}}_\tau(n-1) \\
& - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \bar{\mathcal{M}}_\tau(n) + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \bar{\mathcal{M}}_{\tau'}(n)
\end{aligned} \tag{26b}$$

The second margin,  $\bar{\mathcal{H}}_\tau(\hat{q})$ , is governed by the following Kolmogorov forward equation:

$$\begin{aligned}
(z_I^\tau - \zeta)\hat{q} \frac{d\bar{\mathcal{H}}_\tau(\hat{q})}{d\hat{q}} = & -(\delta_\tau + d_\tau + \eta - z_X^\tau)\bar{\mathcal{H}}_\tau(\hat{q}) + \mu_e \frac{m_\tau}{M_\tau} (1 - \Phi(\hat{q})) \\
& - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \bar{\mathcal{H}}_{\tau'}(\hat{q}) + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{M_{\tau'}}{M_\tau} \bar{\mathcal{H}}_{\tau'}(\hat{q}).
\end{aligned} \tag{27}$$

Lastly, the distribution of firm sizes is some ‘‘convolution’’ of the distribution of the number of establishments per firm and the distribution of establishment sizes and can be derived as follows. Let  $\mathcal{M}_\tau(n, \hat{q})$  be the normalized measure of type- $\tau$  firms with  $n$  establishments and  $q(t) \geq \hat{q}Q(t)$ . For each  $\tau \in \Gamma$ , we have

$$\begin{aligned}
(z_I^\tau - \zeta)\hat{q} \frac{d\mathcal{M}_\tau(1, \hat{q})}{d\hat{q}} = & -(z_X^\tau + \delta_\tau + d_\tau + \eta)\mathcal{M}_\tau(1, \hat{q}) \\
& + 2\delta_\tau\mathcal{M}_\tau(2, \hat{q}) + \mu_e m_\tau (1 - \Phi(\hat{q})) \\
& + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \mathcal{M}_{\tau'}(1, \hat{q}) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \mathcal{M}_\tau(1, \hat{q}).
\end{aligned} \tag{28a}$$

and for  $n > 1$ :

$$\begin{aligned}
(z_I^\tau - \zeta)\hat{q} \frac{d\mathcal{M}_\tau(n, \hat{q})}{d\hat{q}} = & -(n(z_X^\tau + \delta_\tau) + d_\tau + \eta)\mathcal{M}_\tau(n, \hat{q}) \\
& + (n+1)\delta_\tau\mathcal{M}_\tau(n+1, \hat{q}; t) + (n-1)z_X^\tau\mathcal{M}_\tau(n-1, \hat{q}) \\
& + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \mathcal{M}_{\tau'}(n, \hat{q}) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \mathcal{M}_\tau(n, \hat{q}).
\end{aligned} \tag{28b}$$

The detailed derivations of these equations are presented in Appendix C.

### 4.3.2 Distributions for one-type economy

When there is only one firm type, firm growth is governed by three endogenous numbers:  $z_I$ ,  $z_X$ , and  $\mu_e$ . Note that, in this case, our model assumptions imply that for a given firm, the quality (and therefore the size) of each establishment grows at a deterministic rate  $z_I$  that is common across all firms. The average quality  $Q(t)$  grows at the rate  $\zeta$  given in (23). Thus, the quality of the establishments in a firm who started at time  $t_0$  and whose initial draw of the normalized quality is

$\hat{q}Q(t_0)$  can be represented as (denoting it by  $q_{t_0}(t)$ )

$$q_{t_0}(t) = \hat{q}Q(t_0)e^{z_I(t-t_0)} = \hat{q}Q(t)e^{(z_I-\zeta)(t-t_0)}.$$

From the labor demand (7) and the labor market equilibrium condition (13), it is straightforward to show that along the balanced-growth equilibrium, the relative labor demand  $\ell(t)/L(t)$  of a particular establishment with quality  $q_{t_0}(t)$  is equal to  $q_{t_0}(t)/(N(t)Q(t))$ . Therefore, the cross-sectional distribution of establishment size at a given time  $t$  is the same as the distribution of  $\hat{q}e^{(z_I-\zeta)(t-t_0)}$ . Denoting the time- $t$  number of establishments for a firm that starts at time  $t_0$  as  $n_{t_0}(t)$  (note that  $n_{t_0}(t)$  is stochastic as the external innovation is random), the (relative) firm size distribution follows the distribution of  $n_{t_0}(t)\hat{q}e^{(z_I-\zeta)(t-t_0)}/N(t)$ .

Equation (27) becomes

$$(z_I - \zeta)\hat{q}\frac{d\tilde{\mathcal{H}}(\hat{q})}{d\hat{q}} = -(\delta + d + \eta - z_X)\tilde{\mathcal{H}}(\hat{q}) + \mu_e(1 - \Phi(\hat{q}))$$

Let us use the change of variables  $p \equiv \log(\hat{q})$  and  $\tilde{\mathcal{H}}(p) \equiv \tilde{\mathcal{H}}(\exp(p))$  to rewrite this equation as:

$$(z_I - \zeta)\frac{d\tilde{\mathcal{H}}(p)}{dp} = -(\delta + d + \eta - z_X)\tilde{\mathcal{H}}(p) + \mu_e(1 - \Phi(\exp(p))).$$

This is a first-order ODE which has a general solution:

$$\tilde{\mathcal{H}}(p) = e^{\frac{\delta+d+\eta-z_X}{z_I-\zeta}(p-p)}\tilde{\mathcal{H}}(\underline{p}) + \int_{\underline{p}}^p e^{\frac{\delta+d+\eta-z_X}{z_I-\zeta}(\tilde{p}-p)}\frac{\mu_e}{z_I-\zeta}(1 - \Phi(\exp(\tilde{p})))d\tilde{p},$$

for each  $\underline{p}$ . Taking the limit  $\underline{p} \rightarrow -\infty$ , and using (24) to replace  $\mu_e$  with  $\delta + d + \eta - z_X$ , we arrive at:

$$\tilde{\mathcal{H}}(p) = \int_{-\infty}^p e^{\frac{\delta+d+\eta-z_X}{z_I-\zeta}(\tilde{p}-p)}\frac{\delta + d + \eta - z_X}{z_I - \zeta}(1 - \Phi(\exp(\tilde{p})))d\tilde{p}. \quad (29)$$

This expression shows that  $\tilde{\mathcal{H}}(\log x)$  is the complementary cumulative distribution function of a random variable  $\mathbf{X}$  defined by a convolution between a Pareto distribution with scale parameter 1 and tail index  $(\delta + d + \eta - z_X)/(z_I - \zeta)$  and a distribution with cumulative distribution function (CDF)  $\Phi$ . That is,  $\mathbf{X}$  is expressed as  $\mathbf{X} = \mathbf{Y}\mathbf{Z}$ , where  $\mathbf{Y} \sim \text{Pareto}(1, (\delta + d + \eta - z_X)/(z_I - \zeta))$  and  $\mathbf{Z} \sim \Phi$ .

Notice also that, when  $\Phi$  is a log-normal distribution,  $\tilde{\mathcal{H}}$  is a convolution of a Pareto distribution and a log-normal distribution analyzed in Reed (2001), and more recently, Cao and Luo (2017) and Sager and Timoshenko (2018). Therefore, we offer an alternative micro-foundation of this convolution distribution with endogenous establishment growth rate, relative to the micro-foundation in Reed (2001) with exogenous growth rate. Our micro-foundation is also more general because it allows for any distribution of  $\Phi$ , while Reed (2001) only allows  $\Phi$  to be a log-normal distribution.

Using this explicit solution, it is easy to show that when  $\Phi$  has thin right tail, for example when



$\Phi$  is a (left-truncated) log-normal distribution, and  $z_I > \zeta$ ,  $\bar{\mathcal{H}}(p)$  has a Pareto tail with the index given by:<sup>15</sup>

$$\frac{\delta + d + \eta - z_X}{z_I - \zeta}. \quad (30)$$

For the distribution of the number of establishments per firm, (26) becomes

$$0 = -(z_X + \delta + d + \eta)\bar{\mathcal{M}}(1) + 2\delta\bar{\mathcal{M}}(2) + \mu_e$$

and for  $n > 1$ :

$$0 = -(n(z_X + \delta) + d + \eta)\bar{\mathcal{M}}(n) + (n + 1)\delta\bar{\mathcal{M}}(n + 1) + (n - 1)z_X\bar{\mathcal{M}}(n - 1)$$

Luttmer (2011) provides a closed form solution for  $\{\bar{\mathcal{M}}(n)\}_{n=1}^{\infty}$ :

$$\bar{\mathcal{M}}(n) = \frac{1}{n} \frac{\mu_e}{z_X} \sum_{k=0}^{\infty} \frac{1}{\beta_{n+k}} \left( \prod_{m=n}^{n+k} \beta_m \right) \prod_{m=1}^{n+k} \frac{z_X \beta_m}{\delta}, \quad (31)$$

where the sequence  $\{\beta_n\}_{n=0}^{\infty}$  is defined recursively by  $\beta_0 = 0$  and

$$\frac{1}{\beta_{n+1}} = 1 - \frac{z_X \beta_n}{\delta} + \frac{\eta + d + z_X n}{\delta n}.$$

Luttmer (2011) also shows that when  $z_X > \delta$ ,  $\bar{\mathcal{M}}(n)$  has Pareto tail with the index given by

$$\frac{\eta + d}{z_X - \delta}. \quad (32)$$

The following proposition summarizes the last two results.

**Proposition 5** *On a stationary BGP with  $z_I > \zeta$  and  $z_X > \delta$ , the stationary distribution of establishment sizes and the stationary distribution of the number of establishments per firm have Pareto right tail and the tail indices are given by (30) and (32) respectively.*

Now, we analyze the distribution of firm sizes. In a special case where the initial draw satisfies  $\int \hat{q} d\Phi(\hat{q}) = 1$ , (23) implies that  $z_I = \zeta$  holds, and thus the establishment size distribution is identical to the distribution of  $\hat{q}$ . Furthermore, in this case, the distribution of  $\hat{q}$  for a given  $n(t)$  follows the identical distribution  $\Phi(\hat{q})$ . The distribution of  $n(t)$  evolves with the standard birth-death process, as in Klette and Kortum (2004) and Luttmer (2011). Let the number of firms with  $n(t) = n$  in the balanced-growth path be  $\bar{\mathcal{M}}(n)N(t)$ .<sup>16</sup>

The  $N(t)$ -normalized measure of firms with  $n(t) = n$  and  $q(t) \geq \hat{q}Q(t)$ ,  $\mathcal{M}(n, \hat{q})$ , is

$$\mathcal{M}(n, \hat{q}) = \bar{\mathcal{M}}(n)(1 - \Phi(\hat{q})),$$

<sup>15</sup>Using (23) and (24), this tail index is also equal to  $\frac{1}{1 - \int \frac{1}{\hat{q} d\Phi(\hat{q})}}$ .

<sup>16</sup>Note that below we normalize all measures of firms, such as  $\bar{\mathcal{M}}(n)$ , by  $N(t)$ . This is because the number of firms grows at the same rate as  $N(t)$  does. Normalizing by  $N(t)$  implies that  $\sum_n n\bar{\mathcal{M}}(n) = 1$  and the fraction of firms with  $n(t) = n$  can be obtained by calculating  $\bar{\mathcal{M}}(n)/\sum_n \bar{\mathcal{M}}(n)$ .

where  $\bar{\mathcal{M}}(n)$  is given by (31). Once we have  $\mathcal{M}(n, \hat{q})$ , the firm size distribution can be computed easily. The fraction of firms with size  $\mathbf{l}(t) \geq \hat{\mathbf{L}}(t)$ , denoted by  $\mathbf{M}(\hat{\mathbf{l}})$ , can be computed as

$$\mathbf{M}(\hat{\mathbf{l}}) = \frac{\sum_n \mathcal{M}(n, \hat{\mathbf{l}}/n)}{\sum_n \bar{\mathcal{M}}(n)}.$$

To determine the tail index of  $\mathbf{M}(\cdot)$ , we consider the Laplace transformation:<sup>17</sup>

$$\varphi(s) = \int_0^\infty \hat{l}^s \left( -d\mathbf{M}(\hat{l}) \right). \quad (33)$$

Using the expression for  $\mathbf{M}$  above, we rewrite  $\varphi$  as:

$$\begin{aligned} \varphi(s) &= \int_0^\infty \hat{l}^s \frac{-d \sum_n \bar{\mathcal{M}}(n) (1 - \Phi(\hat{l}/n))}{\sum_n \bar{\mathcal{M}}(n)} \\ &= \int_0^\infty \hat{l}^s \frac{\sum_n \bar{\mathcal{M}}(n) d\Phi(\hat{l}/n)}{\sum_n \bar{\mathcal{M}}(n)} \\ &= \int_0^\infty \frac{\sum_n \bar{\mathcal{M}}(n) n^s (\hat{l}/n)^s d\Phi(\hat{l}/n)}{\sum_n \bar{\mathcal{M}}(n)} \\ &= \left\{ \int_0^\infty \hat{l}^s d\Phi(\hat{l}) \right\} \left\{ \frac{\sum_n \bar{\mathcal{M}}(n) n^s}{\sum_n \bar{\mathcal{M}}(n)} \right\} \end{aligned}$$

Assume that the entry distribution  $\Phi$  has thin right tail and using the closed form solution in Luttmer (2011), we can show that

$$\varphi(s) \sim \frac{c}{\frac{\eta + d}{z_X - \delta} - s}$$

as  $s \uparrow (\eta + d)/(z_X - \delta)$ . Therefore, by the Tauberian theorems in Mimica (2016),  $\mathbf{M}$  has right Pareto tail with the tail index given by  $(\eta + d)/(z_X - \delta)$ .

The following proposition summarizes the result.

**Proposition 6** *On a stationary BGP with  $z_I = \zeta$  (or  $z_I < \zeta$ ) and  $z_X > \delta$ , and the distribution of entry sizes  $\Phi$  has thin right tail, firm size distribution has Pareto right tail with the tail index equals to the tail index of the distribution of the number of establishments per firm given by (32).*

Now we consider the more challenging case in which  $z_I > \zeta$ , i.e.  $\int \hat{q} d\Phi(\hat{q}) < 1$ . The system of differential equations (28) for  $\mathcal{M}(n, \hat{q})$  simplifies to:

$$(z_I - \zeta) \hat{q} \frac{d\mathcal{M}(1, \hat{q})}{d\hat{q}} = -(z_X + \delta + d + \eta) \mathcal{M}(1, \hat{q}) + 2\delta \mathcal{M}(2, \hat{q}) + \mu_e (1 - \Phi(\hat{q}))$$

---

<sup>17</sup>After a change of variable  $\hat{l} = \exp(p)$ , the transformation can be re-written in its more familiar form:  $\int_{-\infty}^\infty e^{sp} (-d\mathbf{M}(\exp(p)))$ .

and

$$(z_I - \zeta)\hat{q}\frac{d\mathcal{M}(n, \hat{q})}{d\hat{q}} = -(n(z_X + \delta) + d + \eta)\mathcal{M}(n, \hat{q}) + (n + 1)\delta\mathcal{M}(n + 1, \hat{q}) + (n - 1)z_X\mathcal{M}(n - 1, \hat{q})$$

for  $n > 1$ . Multiplying both sides of these equations with  $\hat{q}^{s-1}$  and integrate from 0 to  $\infty$  then integrating by parts

$$\int_0^\infty \hat{q}^{s-1}\mathcal{M}(n, \hat{q})d\hat{q} = -\frac{1}{s}\int_0^\infty \hat{q}^s d\mathcal{M}(n, \hat{q}),$$

we obtain:

$$-(z_I - \zeta)s\hat{\varphi}(1, s) = -(z_X + \delta + d + \eta)\hat{\varphi}(1, s) - 2\delta\hat{\varphi}(1, s) + \int_0^\infty \hat{q}^{s-1}\mu_e(1 - \Phi(\hat{q}))$$

and

$$-(z_I - \zeta)s\hat{\varphi}(n, s) = -(n(z_X + \delta) + d + \eta)\hat{\varphi}(n, s) + (n + 1)\delta\hat{\varphi}(n + 1, s) + (n - 1)z_X\hat{\varphi}(n - 1, s)$$

for  $n > 1$ , where

$$\hat{\varphi}(n, s) \equiv \int_0^\infty \hat{q}^s(-d\mathcal{M}(n, \hat{q})).$$

For each  $s \geq 0$ , the equations form a system of difference equations and allow us to solve for  $\hat{\varphi}(n, s)$  for all  $n \geq 1$  using the closed form solution from Luttmer (2011) (with  $\eta$  being replaced by  $\eta - (z_I - \zeta)s$ ). We can also show that  $\{\hat{\varphi}(n, s)\}_{n=1}^\infty$  has Pareto tail with the tail index given by

$$\frac{d + \eta - (z_I - \zeta)s}{z_X - \delta}.$$

Now, with the solution for  $\hat{\varphi}(n, s)$ , we can calculate the Laplace transform (33) as follows:

$$\begin{aligned} \varphi(s) &= \int_0^\infty \hat{l}^s \frac{-d \sum_n \mathcal{M}(n, \hat{l}/n)}{\sum_n \bar{\mathcal{M}}(n)} \\ &= \frac{1}{\sum_n \bar{\mathcal{M}}(n)} \sum_n n^s \int_0^\infty (\hat{l}/n)^s (-d\mathcal{M}(n, \hat{l}/n)) \\ &= \frac{\sum_n n^s \hat{\varphi}(n, s)}{\sum_n \bar{\mathcal{M}}(n)}. \end{aligned}$$

Using the tail property of  $\hat{\varphi}(n, s)$ , it is easy to show that  $\varphi(s)$  is finite up to  $s^*$  determined by

$$\frac{d + \eta - (z_I - \zeta)s}{z_X - \delta} = s,$$

or equivalently

$$s^* = \frac{\eta + d}{z_X - \delta + z_I - \zeta}. \quad (34)$$

Therefore  $\mathbf{M}$  has Pareto tail with the tail index given by  $s^*$  (again using the Tauberian theorems

in Mimica (2016)). We summarize these derivations in the following proposition.

**Proposition 7** *On a stationary BGP with  $z_X > \delta$ , firm size distribution has Pareto tail with the tail index given by (34).*

### 4.3.3 Discussion for one-type economy

The model features thick tails for all three distributions of interest (total number of workers in the firm, number of workers per establishment within the firm, and number of establishments within the firm). The main mechanism that generates these tails is firm growth with random exits. As a firm grows by adding establishments and adding workers to those establishments, there is a Poisson arrival of exit shocks ( $\delta$  and  $d$ ) that prevent the firm from continuing to grow. Therefore, the relatively small mass of lucky firms that do not receive an exit shock will continue to grow. The luckiest firms will eventually grow to be very large relative to the average firm and create a thick upper tail of the firm size distribution(s). For the one-type characterization in Section 4.3.2, the expressions for the Pareto tails reflect these forces.

From Proposition 6 and equation (32), we see that the distribution over the number of establishments has a Pareto tail equal to  $(\eta + d)/(z_X - \delta)$ . This expression relates the aggregate inflow of new establishments and outflow of firms through random exit to the net growth rate for establishments ( $z_X - \delta$ ). The distribution's tail is fatter when additional establishments accumulate faster, given random exits induced by establishment and firm shocks  $\delta$  and  $d$ . Using the decomposition of the final output growth rate in equations (23) and (24), we can rewrite the establishment number distribution's Pareto tail as:

$$1 + \frac{\mu_e}{z_X - \delta}.$$

This expression shows that the upper tail of the establishment number distribution is fatter when there is less firm entry and incumbent firms' net growth is faster. Both elements suggest that tail fatness is related to selection on random exits: with fewer firms due to less entry, a fat upper tail only emerges if firms grow fast enough during their finite lifetimes.

Similarly, from equation (30), we see that the distribution over the number of workers per establishment within a firm has a Pareto tail equal to  $[(\eta + d)/(z_I - \zeta)] - [(z_x - \delta)/(z_I - \zeta)]$ . This expression is similar to that for the establishment number distribution. The first term describes a fat upper tail as arising from low churn in terms of establishment inflow and firm death relative to high net growth in establishment size ( $z_I - \zeta$ ), where growth in intensive margin investments to firm quality is large relative to aggregate quality growth  $\zeta$  due to the effect of entrants. The second term is simply an adjustment for the net growth in establishment number. Again using equations (23) and (24), this can be written as:

$$\frac{1}{1 - \int_{\mathcal{N}} \hat{q} d\Phi(\hat{q})}.$$

This version of the expression makes clear that the distribution exhibits a fatter tail when the

average entrant is small relative to the average incumbent, so that  $\int \hat{q} d\Phi(\hat{q})$  is far less than 1. In this case, the firm must grow over time to catch up to the average incumbent firm, during which time random exits pare down the mass of growing firms. The more firms need to catch up to the average incumbent, the more inequality we observe in the distribution.

Finally, we can study the characterization of the overall firm size distribution. From Proposition 8 and equation (34), the Pareto tail of the firm size distribution is equal to  $(\eta + d)/(z_X - \delta + z_I - \zeta)$ , which tells us that a fat upper tail emerges when the firm grows fast – on either margin – relative to the churn of establishment inflows and firm deaths  $(\eta + d)$ . This expression can also be rewritten using equations (23) and (24):

$$\frac{1}{1 - \frac{\mu_e}{z_X - \delta + \mu_e} \int_{\mathcal{N}} \hat{q} d\Phi(\hat{q})}$$

so that the distribution exhibits a fatter upper tail when entrants are small relative to the average incumbent (an effect that derives from the establishment size distribution) and when the entry rate of new firms is low relative to net growth in establishment creation (an effect that derives from the establishment number distribution).

#### 4.3.4 Distributions for two-type economy

Consider an economy with only two type  $\tau \in \{L, H\}$ . To simplify the derivations of stationary distributions, we assume that  $0 = \lambda_{LH} < \lambda_{HL}$ , i.e. the high-type can transition to becoming a low type, but not vice versa.

Using the procedure employed by Luttmer (2011), one can show that, under some parameter restrictions, the stationary distribution of the number of establishments per firm has approximate right Pareto tail with the tail index given by:

$$\min \left\{ \frac{\eta + \lambda_{HL} + d_H}{[z_X^H - \delta_H]_+}, \frac{\eta + d_L}{[z_X^L - \delta_L]_+} \right\}. \quad (35)$$

This corresponds to (32) in the case of single type.

Similar, using the procedure from Gabaix et al. (2016) and Cao and Luo (2017), the Pareto tail index of the distribution of establishment sizes is given by

$$\min \left\{ \frac{\eta + \delta_H + \lambda_{HL} - z_X^H}{[z_I^H - \zeta]_+}, \frac{\eta + \delta_L + d_L - z_X^L}{[z_I^L - \zeta]_+} \right\}. \quad (36)$$

Lastly, assuming that  $z_X^L \leq \delta_L$ , using Laplace transform as in Subsection 4.3.2, we can show that the Pareto tail index of firm size distribution is given by

$$\frac{\eta + \lambda_{HL} + d_H}{z_X^H - \delta_H + z_I^H - \zeta}. \quad (37)$$

## 5 Quantitative analysis

In this section, we estimate our model to quantitatively analyze the pattern of firm growth over the 1995-2015 period. First we develop an estimation strategy for recovering model parameters from firm size, establishment number, and establishment size distributions. Then we confirm that the model is able to match these firm size distributions well. Finally, we estimate the model for both 1995 and 2015 data and show how the model informs us about fundamental economic forces that have changed with average firm size growth over these years.

### 5.1 Construction of data moments

The estimation of the model in this section uses the cross-sectional distribution of establishment size and the number of establishment. We are awaiting the completion of a disclosure avoidance review in order to obtain the data moments that will be used in the model estimation.

In this draft, we take an alternative approach. We assume that the data is drawn from simple parametric distributions that are known to fit the actual U.S. data from past studies. We first estimate these distributions using publicly available data on firm size distributions as well as a subset of the aggregate moments we already disclosed. This procedure provides us the data moments that we will substitute by the moments that directly come from our dataset once the disclosure process is complete. The data moments include the Pareto tail index of the establishment size distribution and distribution of establishment number which cannot be inferred directly from publicly available data. These tail indices are crucial for our estimation procedure.

First, for the distribution the number of establishments for each firm (call it  $\mathbf{N}$ ), we assume that in year  $t$ , the distribution takes the following form:

$$\Pr_t(\mathbf{N} \leq n) = G(\log n; \mu^{ne}, \sigma^{ne}, \lambda_t^{ne}),$$

for  $n = 1, 2, \dots$ , where  $G$  is the CDF of the convolution between a normal distribution and an exponential distribution (see Sager and Timoshenko (2018) for more details on this type of distribution):

$$G(z; \mu, \sigma, \lambda) \equiv \Phi_n \left( \frac{z - \mu}{\sigma} \right) - e^{-\lambda(z - \mu) + \frac{\sigma^2}{2} \lambda^2} \Phi_n \left( \frac{z - \mu}{\sigma} \right),$$

where  $\Phi_n$  is the cdf of the standard normal distribution. This distribution flexibly nests both a normal distribution and an exponential distribution and conveniently allows for a thick right tail. We estimate  $\mu^{ne}$ ,  $\sigma^{ne}$ , and  $\lambda_{2005}^{ne}$  by targeting the mean number of establishment per firm of 1.4163 reported in Figure 6 and the third column of Table 1 published by Bureau of Labor Statistics.<sup>18</sup> The second column of Table 1 gives a coarse characterization of the distribution over the number

---

<sup>18</sup>The numbers are taken from 2005 Q1 Business Employment Dynamics (which is drawn from the Quarterly Census of Employment and Wages, explained in Section 2.1): <https://www.bls.gov/bdm/sizeclassqanda.htm>.

Table 1: Distribution of firms by number of establishments, first quarter 2005.

	Number of firms (thousands)	Data number of firms (% of total)	Synthetic distribution
1 establishment	4,706	95.2	95.20
2 establishments	112	2.3	2.22
3 establishments	36	0.7	0.83
4 establishments	20	0.4	0.43
5-9 establishments	38	0.8	0.75
10 or more establishments	30	0.6	0.56
Mean number of establishments	1.42		1.42

of establishments per firm, which parameters of  $G(\cdot)$  will match. The estimation yields

$$\mu^{ne} = -4.6748,$$

$$\sigma^{ne} = 2.2129,$$

and

$$\lambda_{2005}^{ne} = 1.2111.$$

The resulting synthetic distribution fits the actual U.S. data moments very well, as shown in the comparison of columns 2 and 3 of Table 1.

Table 2: Distribution of establishments by employment

Establishment Size Bin	% of total, 2015		% of total, 1995	
	Data	Synthetic	Data	Synthetic
(a) 1-4	48.94	53.08	49.15	53.78
(b) 5-9	21.11	18.57	22.62	18.82
(c) 10-19	14.26	13.05	13.87	12.91
(d) 20-49	9.80	9.67	8.95	9.27
(e) 50-99	3.27	3.33	3.00	3.10
(f) 100-249	1.86	1.70	1.71	1.56
(g) 250-499	0.47	0.40	0.43	0.37
(h) 500-999	0.18	0.13	0.16	0.13
(i) 1000+	0.10	0.06	0.09	0.06
Mean size	12.40	15.26	12.80	15.03

In our quantitative exercise, we are ultimately interested in comparing firm growth at 1995 and 2015. To construct synthetic distributions for 1995 and 2015, we hold  $\mu^{ne}, \sigma^{ne}$  constant while varying  $\lambda^{ne}$  to match the mean number of establishments (1.29 for 1995 and 1.55 for 2015, as in

Figure 6). The resulting estimates are

$$\lambda_{1995}^{ne} = 1.3060$$

and

$$\lambda_{2015}^{ne} = 1.1606.$$

Note that this procedure recovers tail parameters that decline over time, which implies that there is a growing number of large “superstar firms” between 1995 and 2015.

Using the same published source from the BLS, we also construct a synthetic distribution for the establishment size. Similarly to above, we assume that the distribution also takes the exponentially modified Gaussian form:

$$\Pr_t(\mathbf{1} \leq \ell) = G(\log \ell; \mu_t^e, \sigma_t^e, \lambda_t^e),$$

for establishment sizes  $\ell = 1, 2, \dots$  and in years  $t = 1995, 2015$ . We estimate  $\mu_t^e, \sigma_t^e$ , and  $\lambda_t^e$  by targeting Table 2 published by BLS (second and fourth column) for 1995 and 2015, as well as the mean establishment size reported in Figure 5 of Section 2. The estimation yields

$$\mu_{1995}^e = 0.6634, \quad \sigma_{1995}^e = 1.4881, \quad \lambda_{1995}^e = 1.6796$$

and

$$\mu_{2015}^e = 0.7603, \quad \sigma_{2015}^e = 1.5542, \quad \lambda_{2015}^e = 1.9425.$$

The targeted moments are closely replicated as shown in Table 2 published by BLS (third and fifth column). Notice that the establishment size distribution has a tail parameter that increases over time ( $\lambda_{2015}^e > \lambda_{1995}^e$ ), indicating that skewness in the establishment size distribution has declined over time.

Lastly, we construct a synthetic distribution for the firm size using the exponentially modified Gaussian form:

$$\Pr_t(\mathbf{1} \leq \ell) = G(\log \ell; \mu_t^f, \sigma_t^f, \lambda_t^f),$$

for firm sizes  $\ell = 1, 2, \dots$  and in years  $t = 1995, 2015$ . We estimate  $\mu_t^f, \sigma_t^f$ , and  $\lambda_t^f$  by targeting Table 3 published by BLS (second and fourth column) for 1995 and 2015, as well as the mean firm size reported in Figure 4 of Section 2. The estimation yields

$$\mu_{1995}^f = 0.3881, \quad \sigma_{1995}^f = 1.1383, \quad \lambda_{1995}^f = 1.1385$$

and

$$\mu_{2015}^f = 0.3377, \quad \sigma_{2015}^f = 1.2198, \quad \lambda_{2015}^f = 1.1301.$$

The targeted moments are closely replicated as shown in Table 3 published by BLS (third and fifth column) except for the very large firms. The tail parameters stay relatively stable, only declines very slightly, from 1995 to 2015.



Table 3: Distribution of firms by employment

Establishment Size Bin	% of total, 2015		% of total, 1995	
	Data	Synthetic	Data	Synthetic
(a) 1-4	55.918	56.627	55.017	55.879
(b) 5-9	20.044	19.356	21.324	20.351
(c) 10-19	11.998	11.840	12.144	12.080
(d) 20-49	7.605	7.629	7.337	7.456
(e) 50-99	2.342	2.455	2.261	2.311
(f) 100-249	1.294	1.350	1.207	1.247
(g) 250-499	0.392	0.403	0.345	0.368
(h) 500-999	0.190	0.184	0.163	0.166
(i) 1000-2499	0.118	0.099	0.107	0.089
(j) 2499-4999	0.044	0.029	0.041	0.026
(k) 5000-9999	0.024	0.013	0.023	0.012
(l) 10000+	0.028	0.014	0.024	0.010
Mean size	25.50	26.13	23.0	23.6

## 5.2 Model Estimation

First, we describe our two-step estimation procedure. Then we present the estimates and associated quantitative results.

### 5.2.1 Computing equilibrium

We can compute the equilibrium as follows. First, make a guess on  $g$ . Once  $g$  is known, we can compute  $r$  by  $r = \rho + \sigma g$  and look for  $\bar{w}$  that satisfies the free-entry condition (11) by solving (9) for given  $r$  and  $\bar{w}$ . Then, we compare the initial guess with the  $g$  calculated from the general equilibrium. In particular, the set of equations (17), (19), (21), and (22) can be used for pinning down  $\mu_e$ ,  $g$ ,  $M_\tau$ , and  $s_\tau$  ( $2T + 2$  equations and  $2T + 2$  unknowns). If this  $g$  is different from the initial guess, adjust  $g$  until we find the fixed point.

When  $\alpha = 1$ , the algorithm is slightly different. Because we can solve for  $g$  from (12) as a function of parameters (exogenous growth), we can obtain  $r$  without making any guess. After solving the HJB equations and finding  $\bar{w}$  that satisfies free-entry condition, we can use (21) and (22) to obtain  $\mu_e$  and  $s_\tau$ . Given these,  $\eta$  and  $M_\tau$  can be solved from (17) and (18).

### 5.2.2 Overview of the estimation procedure

Although the analytical characterization of the model with one type provided us useful insights, Appendix D.1 shows that at least two types are needed to match both the extensive and intensive margin distributions. As such, we estimate a simple version of the model with two types:  $\Gamma = \{H, L\}$ .  $H$ -type firms have lower cost of external investment and expand their number of es-

establishments faster.  $H$ -type firms transition to  $L$ -type firms at the rate  $\lambda_{HL} > 0$  while  $L$ -type firms do not transition to  $H$ -type firms, i.e.  $\lambda_{LH} = 0$ . Thus the  $L$ -type is the absorbing state. These are similar to the assumptions made in Luttmer (2011) (but without differentiating the extensive versus intensive margins as we do here). These assumptions allow us to easily obtain closed-form solutions for the distribution of the number of establishments per firm and we use these closed-form solutions to estimate the model.

We estimate the model in two steps. In Step 1, we estimate  $(z_X^H, z_X^L, \lambda_{HL}, \mu_e, m_H, m_L)$  using the moments related to the number of establishments per firm; and then we estimate  $(z_I^H, z_I^L, \Phi(\cdot))$  using the moments related to the number of employees per establishments. In Step 2, we assume functional forms for the cost functions  $h_X, h_I$  and estimate the parameters of these functions using the estimates from Step 1.

**Step 1a (Number of establishments per firm):** In this step, we choose

$$(z_X^H, z_X^L, \lambda_{HL}, \mu_e, m_H, m_L)$$

parameters to target (i) percentiles of the distribution over the number of establishments per firm, (ii) the slope of the upper tail of the distribution, and (iii) the growth rate of the number of the establishments  $\eta \approx 1\%$ . With two types, (16) becomes

$$\begin{aligned} \eta &= z_X^H - \delta_H - d_H + \mu_e \frac{m_H}{M_H} - \lambda_{HL} \\ &= z_X^L - \delta_L - d_L + \mu_e \frac{m_L}{M_L} + \lambda_{HL} \frac{M_H}{M_L}. \end{aligned}$$

Together with  $M_H + M_L = 1$ , the last equality gives us a unique solution for  $M_H, M_L$ . In particular,

$$M_L = \frac{-a_1 - \sqrt{a_1^2 - 4a_0a_2}}{2a_0}$$

where

$$\begin{aligned} a_0 &= (z_X^H - \delta_H - d_H) - (z_X^L - \delta_L - d_L) \\ a_1 &= -(\mu_e(m_H + m_L) + \lambda_{HL}) + ((z_X^L - \delta_L - d_L) - (z_X^H - \delta_H - d_H)) \\ a_2 &= \mu_e m_L + \lambda_{HL}, \end{aligned}$$

and  $M_H = 1 - M_L$ . We then obtain  $\eta$  from either equation.

Having  $\eta$ , we can use formula (35) to calculate the Pareto tail index of the distribution of establishment numbers, i.e., the slope of the distribution's right tail.

**Step 1b (Establishment size):** In this step, we target the distribution of the number of employees per establishments as well as the average growth rate of establishment sizes,  $\zeta$ . The tail index of the distribution of establishment sizes in the model is given by (36). We use this formula to

target the tail index from the empirical distribution of establishment sizes. In addition, we assume that  $\Phi$  follows a log-normal distribution with mean  $\varrho$  and variance  $\varsigma^2$  as in (40) and estimate these parameters to target percentiles of the establishment size distribution.

**Step 2 (Recovering endogenous variables):** In this step we use the estimates from Step 1a and Step 1b, including  $z_i^\tau, \tau \in \{H, L\}, i \in \{X, I\}$ , to quantify the remaining model outcomes and allocations. To execute this step, we must parameterize the innovation cost functions. We assume that the innovation cost functions are iso-elastic of the form  $h_i^\tau(z) \equiv \chi_i^\tau z^\psi$ , for  $\tau \in \Gamma, i \in \{X, I\}$ , and where  $\psi > 0$ . The first order condition in (9) implies  $\psi \chi_i^\tau (z_i^\tau)^{\psi-1} = v_\tau$ , and hence

$$-h_i^\tau(z_i^\tau) + z_i^\tau v_\tau = \left(1 - \frac{1}{\psi}\right) z_i^\tau v_\tau.$$

Substituting this expression in (9) and re-arranging, we arrive at:

$$\underbrace{\begin{bmatrix} A_{11} & -\lambda_{HL} \\ 0 & A_{22} \end{bmatrix}}_{\equiv \mathbf{A}} \begin{bmatrix} v_H \\ v_L \end{bmatrix} = \bar{\pi} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where

$$A_{11} = r - \left(1 - \frac{1}{\psi}\right) (z_X^H + z_I^H) + \delta_H + d_H + \lambda_{HL}$$

and

$$A_{22} = r - \left(1 - \frac{1}{\psi}\right) (z_X^L + z_I^L) + \delta_L + d_L + \lambda_{HL}.$$

From the estimates in Step 1a and Step 1b, all the elements of matrix  $\mathbf{A}$  are known, including  $r = \rho + \sigma g$  and  $g = (\eta + \zeta)$ . Therefore we can then solve for  $v_H, v_L$  as functions of  $\bar{\pi}$ :

$$\begin{bmatrix} v_H \\ v_L \end{bmatrix} = \bar{\pi} \mathbf{A}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Now, combining this result with equations (10) and (11), we obtain

$$\phi = \bar{\pi} \begin{bmatrix} m_H \\ m_L \end{bmatrix}' \mathbf{A}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \exp\left(-\varrho + \frac{\varsigma^2}{2}\right).$$

In other words,  $\bar{\pi}$  is uniquely determined as a function of  $\phi$ .

Lastly, we choose  $\phi$  to match the empirical investment-to-output ratio. On a balanced growth path, we can decompose  $Y(t)$  into  $C(t), R(t), K(t)$  and obtain the following expression for the investment-to-output ratio:

$$\frac{K(t) + R(t)}{Y(t)} = \left(\mu_e \phi + \sum_{\tau} s_{\tau} (h_X^{\tau}(z_X^{\tau}) + h_I^{\tau}(z_I^{\tau}))\right) \left(\frac{\bar{w}(t)}{1 - \beta}\right)^{\frac{1-\beta}{\beta}}. \quad (38)$$

The derivation of this formula is presented in Appendix B.3.

### 5.2.3 Parameter values and estimates

A set of parameters are assigned in advance of estimation, while the remaining parameters are estimated to match empirical moments of the establishment number and size distribution. Table 4 summarizes parameter estimates and targets.

For preferences, for simplicity we assume log utility,  $\sigma = 1$ , and a standard discount rate of  $\rho = 0.01$ . We choose the elasticity of demand for the final good producer to  $\beta = 0.091$ , which is consistent with a markup of 10% (Basu and Fernald (1995)). We assume quadratic innovation costs by setting the elasticity to  $\psi = 2$  as a baseline assumption. With respect to firm and establishment exit rates, we set  $d_H = 0\%$  and  $d_L = 2\%$  at the firm level and  $\delta_H = \delta_L = 12\%$  in 1995 and  $\delta_H = \delta_L = 10\%$  in 2015. These values amount to around 3% quarterly exit rate for establishments and 0.1% quarterly (exogenous) exit rate for firm.<sup>19</sup> Lastly we set population growth rate to the post 1960 average of  $\gamma = 0.011$  and assume purely exogenous growth  $\alpha = 1$  as a baseline assumption.

The remaining parameters are estimated to match empirical moments. We choose an entry cost,  $\phi$ , to match total investment as 10% of total output. We choose  $(\chi_L^I, \chi_H^I, \varrho, \varsigma)$  to match the CDF of the establishment size distribution, where we assume  $\Phi(\cdot)$  is a lognormal distribution with mean  $\varrho$  and standard deviation  $\varsigma$ . We choose  $(\chi_L^X, \chi_H^X, \mu_E, \lambda_{HL}, m_H)$  to match the CDF of the establishment number distribution. We additionally target  $\eta = 0.01$  and  $\zeta = 0.01$  to imply a growth rate of final output of 2%. Finally, we estimate parameters twice—once for 1995 moments of the establishment size and number distributions, and once for 2015 empirical moments.

Figure 16 shows that the distribution of the number of establishments per firm matches the empirical distributions from the (synthetic) 1995 and 2015 data very well. Furthermore, Figure 17 shows that the distribution of the number of employee per establishment closely matches the empirical distribution from the (synthetic) 1995 and 2015 data. The success of the model along these dimensions has to do with the existence of fat tails in the data. The model generates fat tailed distributions endogenously, and therefore the estimation procedure is selecting parameters to match the general slopes of the Pareto tails in the data. Table 5 shows that the model closely matches the remaining empirical targets including Pareto tail estimates computed from the establishment size and number distributions.

Lastly, we can use the estimates given in Table 4 and formula (37) to derive the Pareto tail index of firm size distribution implied by the model. The estimated model produces the tail index in the ballpark of the one obtained from the estimates for synthetic firm size data ( $\lambda^f$ ) in Subsection 5.1. In particular, the tail index for firm size distributions is closer to the tail index for the distribution of number of establishment and to the tail index for establishment size distribution.

---

<sup>19</sup>These parameters can potentially be directly estimated from the our data source to vary by type, which will be the subject of future work.

Table 4: Parameter Values and Targets

Concept	Parameter	Value	Target/Source
<b><i>Parameters Set in Advance</i></b>			
Elasticity of Demand	$\beta$	$1 - (1/1.10)$	10% Markup
Intertemporal Elasticity	$\sigma$	1	Log utility
Discount Rate	$\rho$	0.01	Standard value
Population Growth Rate	$\gamma$	0.011	Census Bureau
Firm Exit Rates	$d_L, d_H$	0.4%, 0%	BLS
Establishment Exit Rates (1995)	$\delta_L, \delta_H$	12%, 12%	BLS
Establishment Exit Rates (2015)	$\delta_L, \delta_H$	10%, 10%	BLS
Innovation Cost	$\psi$	2	Quadratic baseline
Semi-Endogenous Growth	$\alpha$	1	Exogenous baseline
<b><i>Estimated Parameter Targets</i></b>			
Entry Cost	$\phi$		$\frac{Investment}{Output} = 10\%$ , NIPA
Extensive Margin Costs	$\chi_X^H, \chi_X^L$		Establishment number distribution
Intensive Margin Costs	$\chi_X^H, \chi_X^L$		Establishment size distribution
Growth Types	$\lambda_{HL}, m_H$		Establishment number/size right tail
Entrant Size	$\varrho, \varsigma$		Establishment size left tail

Table 5: Moment Fitness

Moments	Data	Model	Data	Model
	1995		2015	
<b><i>Step 1a: Establishment Number Distribution</i></b>				
Establishment flow rate ( $\eta$ )	0.01	0.01	0.01	0.01
Pareto tail index	1.3060	1.3060	1.1606	1.1606
<b><i>Step 1b: Establishment Size Distribution</i></b>				
Average growth rate ( $\zeta$ )	0.01	0.0101	0.01	0.0105
Pareto tail index	1.6796	1.8237	1.9425	2.0620

Notes: Estimates using Least-Square Minimization.

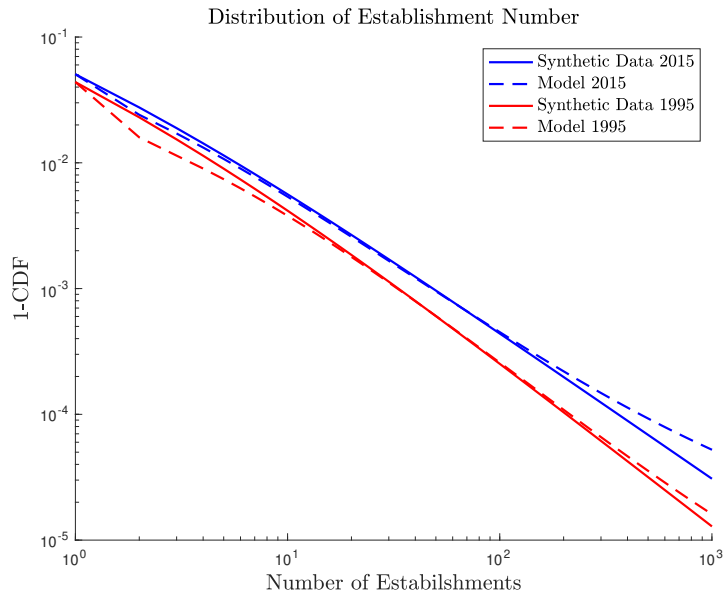


Figure 16: Distribution of number of establishment per firm, Data and Model

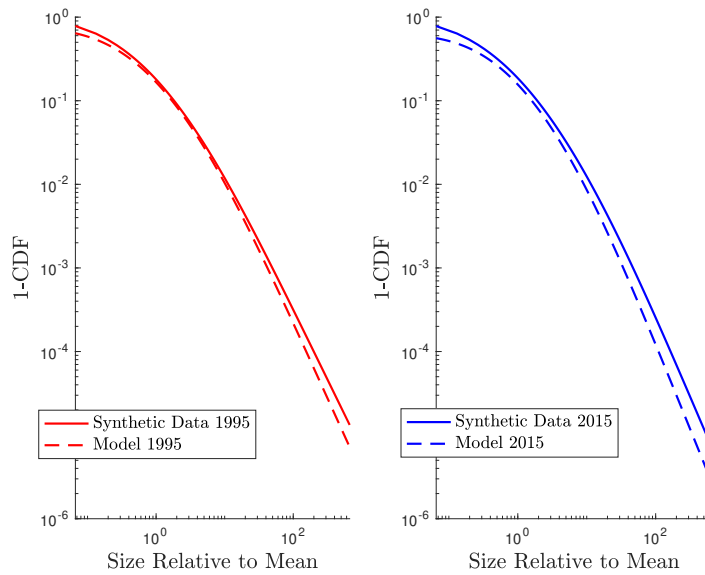


Figure 17: Distribution of number of employee per establishment, Data and Model

Table 6: Parameter Estimates and Model Outcomes, 1995 versus 2005

Parameter	Description	Value (1995)	Value (2015)
<b><i>Innovation Investments</i></b>			
$z_X^H$	<i>H</i> -type external innovation	0.1705	0.2769
$z_X^L$	<i>L</i> -type external innovation	0.0054	0.0028
$z_I^H$	<i>H</i> -type internal innovation	0.0000	0.0000
$z_I^L$	<i>L</i> -type internal innovation	0.0826	0.0652
<b><i>Innovation Costs</i></b>			
$\chi_X^H$	<i>H</i> -type external innovation cost	1.7680	1.0873
$\chi_X^L$	<i>L</i> -type external innovation cost	23.9040	51.5787
$\chi_I^H$	<i>H</i> -type internal innovation cost	$\infty$	$\infty$
$\chi_I^L$	<i>L</i> -type internal innovation cost	1.5610	2.2048
<b><i>Firm Entry</i></b>			
$\mu_E$	Entry rate	0.0999	0.0709
$\phi$	Entry Fixed-Cost	0.0997	0.1428
$\int \hat{q} d\Phi(\hat{q})$	Entrant size relative to mean	0.40	0.47
$\varrho$	Mean of $\Phi(\cdot)$	-1.6917	-1.4393
$\varsigma$	Standard deviation of $\Phi(\cdot)$	1.4153	1.3578
<b><i>Firm Types</i></b>			
$\lambda_{HL}$	<i>H</i> to <i>L</i> transition rate	0.0583	0.1968
$m_H$	Fraction of <i>H</i> -type at entry	0.0255	0.0566
$m_L$	Fraction of <i>L</i> -type at entry	0.9745	0.9434

Notes: Estimates using Least-Square Minimization

### 5.3 Quantitative results

In this section, we compare the results of model estimation from 1995 to that in 2005. Differences in model outcomes over time inform us about the underlying economic mechanisms that generate the observed changes in the distributions over number of establishments within a firm and over average establishment size with a firm. We view this exercise as a growth decomposition of the data through the lens of our theory.

To begin the comparison of 1995 to 2015 outcomes, Table 6 shows that the high-type firm's extensive margin investment rate increased from 17% to 28% as their investment cost coefficients decreased from 1.77 to 1.09. This is an unsurprising result, as theory suggests that the growth in the concentration of firms in the upper tail of the establishment number is partially driven by increased extensive margin investment. Yet, the model simultaneously uncovers that the high-type firm's intensive margin investment rate was zero in both 1995 and 2015, which is driven by very

high investment costs in the model. From the perspective of the model, if an entering firm draws the  $H$ -type, it has an incentive to invest entirely on the extensive margin to take advantage of low R&D cost. The firm invests in intensive margin innovation after it receives the transition shock  $\lambda_{HL}$ . The  $L$ -type firms invest little along the extensive margin (0.5% in 1995) and more than  $H$ -type firms along the intensive margin (8% in 1995). Furthermore, as the establishment size distribution displays a thinner upper tail over time and theory predicts the establishment size distribution's upper tail is partially driven by intensive margin investment rates, we see that  $L$ -type firms invest less in 2015 (6%) than in 1995. Finally, the model shows that  $L$ -type firms decreased their extensive margin investment rate from 0.5% in 1995 to 0.3% in 2015.

The model finds that more entrants are high-growth firms (3% in 1995 and 6% in 2015), yet the average duration of their high growth has shortened from 17 years ( $= 1/0.0583$ ) in 1995 to 5 years ( $= 1/0.1968$ ) in 2015. As such, high-type firms have an increasing incentive to heavily innovate before becoming a  $L$ -type. Given the model's random exit mechanism that underlies the emergence of fat tailed distributions, plus the  $H$ -type firms' incentive to innovate along the extensive margin, fewer firms grow to be increasingly large and hence there is an increasing concentration of large multi-establishment firms in the model.

With respect to entry, the model recovers a declining entry rate of 10% in 1995 to 7% in 2015, driven by an increase in the fixed cost of entry. A declining entry rate is consistent with recent empirical evidence in Decker et al. (2014). The model also recovers an increase in the average initial size of entrant firms relative to the average incumbent, from 40% in 1995 to 47% in 2015. This is consistent with theory, which tells us that an increase in the size of entrants,  $\int \hat{q} d\Phi(\hat{q})$ , leads to a thinner upper tail of the establishment size distribution.

## 6 Conclusion

In this paper, we decomposed firm growth into two margins: an extensive margin of building new establishments and an intensive margin of adding workers to existing establishments. We documented the patterns of extensive and intensive margin firm growth in the U.S. from 1995-2015 and found that U.S. growth is predominantly generated by the addition of new establishments in very large firms. We developed a model of firm growth that incorporates both the extensive and intensive margins as separate types of firm innovations and showed the model can generate a fat tail of large firms – both in terms of number of establishments and number of workers.

We estimate model parameters for 1995 and 2015, and use the model to interpret the increase in the firm size we observe in the data by fundamental economic changes. We found that three factors are important in characterizing the changes in firm dynamics during that period. First, the cost for the type of innovation that leads to building new establishments fell for firms that are innovative in that dimension. Second, the duration of a firm being highly innovative in such innovations became shorter. Third, the firm entry cost increased.

An important future research is why these three changes occurred during the recent time period.



The first fact accords well with the anecdotes that it became easier to find new locations for stores and restaurants due to increasing availability of data. The second fact may also be a reflection of the faster information flow—a successful business model may become obsolete with a faster cycle when it is easily imitated. The third is consistent with the recent literature on increased concentration and reduced dynamism of the U.S. economy.

## References

- Acemoglu, D. and D. Cao (2015). Innovation by entrants and incumbents. *Journal of Economic Theory* 157, 255 – 294.
- Akcigit, U. and W. R. Kerr (2016). Growth through heterogeneous innovations. Working Paper 16443, National Bureau of Economic Research.
- Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2017). The Fall of the Labor Share and the Rise of Superstar Firms. NBER Working Papers 23396, National Bureau of Economic Research, Inc.
- Cao, D. and W. Luo (2017). Persistent heterogeneous returns and top end wealth inequality. *Review of Economic Dynamics* 26, 301 – 326.
- Chodorow-Reich, G. (2014). The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-9 Financial Crisis. *The Quarterly Journal of Economics* 129(1), 1–59.
- Chodorow-Reich, G. and J. Wieland (2017). Secular Labor Reallocation and Business Cycles. Working Paper 313261, Harvard University OpenScholar.
- Choi, E. J. and J. R. Spletzer (2012, March). The declining average size of establishments: evidence and explanations. *Monthly Labor Review*, 50 – 65.
- De Loecker, J. and J. Eeckhout (2017). The rise of market power and the macroeconomic implications. Working Paper 23687, National Bureau of Economic Research.
- Decker, R., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2014). The role of entrepreneurship in u.s. job creation and economic dynamism. *Journal of Economic Perspectives* 28, 3 – 24.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica* 79(3), 733–772.
- Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica* 84(6), 2071–2111.
- Hathaway, I. and R. E. Litan (2014). Declining business dynamism in the united states: A look at states and metros. Brookings institute report.
- Kehrig, M. and N. Vincent (2017). Do firms mitigate or magnify capital misallocation? evidence from plant-level data. *Working Paper*.
- Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy* 112(5), 986–1018.
- Lentz, R. and D. T. Mortensen (2008). An empirical model of growth through product innovation. *Econometrica* 76, 1317–1373.

- Luttmer, E. G. J. (2011). On the mechanics of firm growth. *The Review of Economic Studies* 78(3), 1042–1068.
- Mimica, A. (2016). Exponential decay of measures and tauberian theorems. *Journal of Mathematical Analysis and Applications* 440(1), 266 – 285.
- Mukoyama, T. and S. Osotimehin (2016). Barriers to reallocation and economic growth: the effects of firing costs. University of Virginia Working Paper.
- Ratner, D. (2013). Unemployment insurance experience rating and labor market dynamics. Finance and Economics Discussion Series 2013-86, Board of Governors of the Federal Reserve System.
- Reed, W. J. (2001). The pareto, zipf and other power laws. *Economics Letters* 74(1), 15 – 19.
- Sager, E. and O. A. Timoshenko (2018). The double emg distribution and trade elasticities. Working paper.
- Siemer, M. (2014). Firm entry and employment dynamics in the great recession. Finance and Economics Discussion Series 2014-56, Board of Governors of the Federal Reserve System.
- Xi, X. (2016). Multi-establishment firms, misallocation and productivity. Working paper.

# Appendix

## A Data

This data appendix describes the Quarterly Census of Employment and Wages (QCEW) and draws heavily from the BLS Handbook of Methods.<sup>20</sup>

### A.1 Definitions

The Quarterly Census of Employment and Wages (QCEW) is a count of employment and wages obtained from quarterly reports filed by almost every employer in the U.S., Puerto Rico and the U.S. Virgin Islands, for the purpose of administering state unemployment insurance programs. These reports are compiled by the Bureau of Labor Statistics (BLS) and supplemented with the Annual Refiling Survey and the Multiple Worksite Report for the purpose of validation and accuracy. The reports include an establishment's monthly employment level upon the twelfth of each month and counts any employed worker, whether their position is full time, part time, permanent or temporary. Counted employees include most corporate officials, all executives, all supervisory personnel, all professionals, all clerical workers, many farmworkers, all wage earners and all piece workers. Employees are counted if on paid sick leave, paid holiday or paid vacation. Employees are not counted if they did not earn wages during the pay period covering the 12th of the month, because of work stoppages, temporary layoffs, illness, or unpaid vacations. The QCEW does not count proprietors, the unincorporated self-employed, unpaid family members, certain farm and domestic workers that are exempt from reporting employment data, railroad workers covered by the railroad unemployment insurance system, all members of the Armed Forces, and most student workers at schools. If a worker holds multiple jobs across multiple firms, then that worker may be counted more than once in the QCEW.

### A.2 Sample

Our sample includes each month from 1990 to 2016 and covers 38 states: Alaska, Alabama, Arkansas, Arizona, California, Colorado, Connecticut, Delaware, Georgia, Hawaii, Iowa, Idaho, Indiana, Kansas, Louisiana, Maryland, Maine, Minnesota, Montana, North Dakota, New Jersey, New Mexico, Nevada, Ohio, Oklahoma, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Vermont, Washington, West Virginia, as well as the District of Columbia, Puerto Rico and the U.S. Virgin Islands.

### A.3 Data Cleaning and Variable Construction

To conform to official statistics, we clean the data in accordance with BLS procedure. First, while the QCEW contains monthly data as of the 12th of each month, we follow BLS convention by

---

<sup>20</sup>See <https://www.bls.gov/opub/hom/cew/home.htm> for the complete BLS Handbook of Methods.

only using data from the final month within a quarter. As a result, our sample does not capture establishments that enter and exit within the same quarter. We additionally exclude firms from calculations in a given quarter if the absolute change in employment from the previous quarter exceeds 10 times the average employment between the two quarters. Statistics within this paper are not sensitive to the choice of multiple being 10.

We construct firms as the summation over employment identification numbers (EINs). Firm-level employment is the sum of all employment in establishments associated with the same EIN and the number of establishments within a firm as the number of establishments that report using a common EIN. To classify a firm's industry, we assign to a firm the average self-reported, 6-digit NAICS code of its establishments so that the firm is classified in the same way as its establishments are on average.

A firm's entry date is measured as the date at which the QCEW records a non-zero number of workers associated with a particular EIN after four consecutive quarters of recording zero workers. A firm's exit date is measured as the last date at which the QCEW records a non-zero number of workers associated with a particular EIN prior to four consecutive quarters of recording zero workers. A firm's age is measured by tracking firms after entering. Upon entry, the firm is assigned an age of 1 quarter and the firm's age is incremented by 1 quarter for each period that it does not exit.

## B Derivations

### B.1 Derivation of (1)

$$\begin{aligned} \log(Z) &= \log(E[XY|XY = Z]) \\ &= \log(E[X|XY = Z]) + \log(E[Y|XY = Z]) + \log\left(\frac{E[XY|XY = Z]}{E[X|XY = Z]E[Y|XY = Z]}\right). \end{aligned}$$

Call the final term as  $\Omega$ . Because  $E[XY|XY = Z] = Z$  and  $E[X|XY = Z]E[Y|XY = Z] = E[X|XY = Z]E[Z/X|XY = Z] = E[X|XY = Z]E[1/X|XY = Z]Z$ . From Jensen's inequality,

$$E[X|XY = Z]E\left[\frac{1}{X}\middle|XY = Z\right] \geq E[X|XY = Z]\frac{1}{E[X|XY = Z]} = 1,$$

and thus  $\Omega \leq 0$  and the equality holds when  $\text{var}[X|XY = Z] = 0$ .

### B.2 Derivation of (22)

From the definition of  $Q_\tau(t)$ ,

$$\frac{\dot{Q}_\tau(t)}{Q_\tau(t)} = -\frac{\dot{N}_\tau(t)}{N_\tau(t)} + \frac{d \int_{\mathcal{N}_\tau(t)} q_j(t) dj / dt}{\int_{\mathcal{N}_\tau(t)} q_j(t) dj}. \quad (39)$$

The first term of the right-hand side is  $-\eta$ . To compute the second term, consider a discrete time interval  $\Delta t > 0$ , compute  $(\int_{\mathcal{N}_\tau(t+\Delta t)} q_j(t+\Delta t) dj - \int_{\mathcal{N}_\tau(t)} q_j(t) dj) / \Delta t$  and set  $\Delta t \rightarrow 0$ . Note that the denominator of the second term is equal to  $Q(t)N(t)$ . Because

$$\begin{aligned} & \int_{\mathcal{N}_\tau(t+\Delta t)} q_j(t+\Delta t) dj - \int_{\mathcal{N}_\tau(t)} q_j(t) dj \\ &= [z_I^T + z_X^T] \Delta t Q_\tau(t) N_\tau(t) - (\delta_\tau + d_\tau) \Delta t Q_\tau(t) N_\tau(t) + \mu_e \Delta t m_\tau Q(t) N(t) \int \hat{q} d\Phi(\hat{q}) \\ & \quad - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \Delta t Q_\tau(t) N_\tau(t) + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \Delta t Q_{\tau'}(t) N_{\tau'}(t) + o(\Delta t), \end{aligned}$$

where the first term is the additional quality by internal and external innovation, the second term is the lost quality by exit, the third term is the gain from entry, and the fourth and the fifth terms are the loss and gain from the transitions of firm types. (The higher-order terms are omitted as  $o(\Delta t)$ .) Dividing by  $\Delta t$  and taking  $\Delta t \rightarrow 0$ ,

$$\begin{aligned} & \frac{d \int_{\mathcal{N}_\tau(t)} q_j(t) dj}{dt} \\ &= Q_\tau(t) N_\tau(t) \left( z_I^T + z_X^T - (\delta_\tau + d_\tau) + \mu_e m_\tau \frac{Q(t) N(t)}{Q_\tau(t) N_\tau(t)} \int \hat{q} d\Phi(\hat{q}) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{Q_{\tau'}(t) N_{\tau'}(t)}{Q_\tau(t) N_\tau(t)} \right). \end{aligned}$$

Therefore, (39) can be rewritten as

$$\zeta = -\eta + z_I^T + z_X^T - (\delta_\tau + d_\tau) + \mu_e \frac{m_\tau}{M_\tau} \frac{Q(t)}{Q_\tau(t)} \int \hat{q} d\Phi(\hat{q}) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \frac{Q_{\tau'}(t) M_{\tau'}}{Q_\tau(t) M_\tau}$$

Using the definition of  $s_\tau$  and  $g = \eta + \zeta$ , we can obtain (22).

### B.3 Accounting for Aggregate Resources

In this subsection, we decompose  $Y(t)$  into  $C(t), R(t), K(t)$  on a BGP.

Plugging the optimal choice for  $x_j(t)$ , (4), in the aggregate production function (3), we obtain

$$\begin{aligned} Y(t) &= \left( \int_{\mathcal{N}(t)} q_j(t)^\beta x_j(t)^{1-\beta} dj \right)^{\frac{1}{1-\beta}} \\ &= \left( \int_{\mathcal{N}(t)} q_j(t)^\beta \left[ (1-\beta)^{\frac{1}{\beta}} \left( \frac{w(t)}{Z(t)} \right)^{-\frac{1}{\beta}} Y(t) q_j(t) \right]^{1-\beta} dj \right)^{\frac{1}{1-\beta}} \\ &= (1-\beta)^{\frac{1}{\beta}} \left( \frac{w(t)}{Z(t)} \right)^{-\frac{1}{\beta}} Y(t) \left( \int_{\mathcal{N}(t)} q_j(t) dj \right)^{\frac{1}{1-\beta}} \\ &= \left( \frac{w(t)}{(1-\beta)Z(t)} \right)^{-\frac{1}{\beta}} Y(t) (N(t)Q(t))^{\frac{1}{1-\beta}}, \end{aligned}$$

where the last equality uses the definition of  $Q(t)$ .

Simplifying  $Y(t)$  from both sides, we arrive at:

$$1 = \left( \frac{w(t)}{(1-\beta)Z(t)} \right)^{-\frac{1}{\beta}} (N(t)Q(t))^{\frac{1}{1-\beta}}.$$

Combining this identity with the labor market clearing condition (14) yields

$$Y(t) = L(t)Z(t) (N(t)Q(t))^{\frac{\beta}{1-\beta}}.$$

We then use (14) to replace  $L(t)Z(t)$  with  $\left(\frac{\bar{w}(t)}{1-\beta}\right)^{-\frac{1}{\beta}} N(t)Q(t)Y(t)^{-\frac{\beta}{1-\beta}}$ . Thus

$$Y(t) = \left( \frac{\bar{w}(t)}{1-\beta} \right)^{-\frac{1-\beta}{\beta}} N(t)Q(t).$$

On a BGP,  $K(t) = \mu_e N(t)\phi Q(t)$ . Therefore

$$\frac{K(t)}{Y(t)} = \mu_e \phi \left( \frac{\bar{w}(t)}{1-\beta} \right)^{\frac{1-\beta}{\beta}}$$

The aggregate cost of intensive and extensive margin investment is given by:

$$\begin{aligned} R(t) &= \sum_{\tau} (h_X^{\tau}(z_X^{\tau}) + h_I^{\tau}(z_I^{\tau})) \int_{\mathcal{N}_{\tau}(t)} q_j(t) dj \\ &= \sum_{\tau} (h_X^{\tau}(z_X^{\tau}) + h_I^{\tau}(z_I^{\tau})) N_{\tau}(t) Q_{\tau}(t) \\ &= N(t)Q(t) \sum_{\tau} (h_X^{\tau}(z_X^{\tau}) + h_I^{\tau}(z_I^{\tau})) s_{\tau}, \end{aligned}$$

where  $s_{\tau}$  is defined in (20). As a result

$$\frac{R(t)}{Y(t)} = \left( \sum_{\tau} (h_X^{\tau}(z_X^{\tau}) + h_I^{\tau}(z_I^{\tau})) s_{\tau} \right) \left( \frac{\bar{w}(t)}{1-\beta} \right)^{\frac{1-\beta}{\beta}}.$$

Combining the expressions for the ratios  $K/Y$  and  $R/Y$ , we obtain the fraction of total investment over output in (38).

## C Distributional analyses

### C.1 Derivations of Kolmogorov equations

First notice that  $\bar{\mathcal{M}}_{\tau}(n) = \mathcal{M}_{\tau}(n, 0)$ , therefore (26) is a special case of (28) with  $\hat{q} = 0$ . To derive the latter, let  $\hat{\mathcal{M}}_{n,t,q}$  denote the measure of firms with  $n$  establishments and with each establishment having quality of at least  $\hat{q}Q_t$ . Using standard continuous time manipulation of Poisson processes,

we have

$$\begin{aligned}
\hat{\mathcal{M}}_\tau(1, \hat{q}; t + \Delta t) &= \hat{\mathcal{M}}_\tau\left(1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) - (z_X^\tau + \delta_\tau + d_\tau)\Delta t \hat{\mathcal{M}}_\tau\left(1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) \\
&\quad + 2\delta_\tau \Delta t \hat{\mathcal{M}}_\tau\left(2, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) + \mu_e m_\tau \Delta t N_t \left(1 - \Phi\left(\frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}\right)\right) \\
&\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \Delta t \hat{\mathcal{M}}_{\tau'}\left(1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \Delta t \hat{\mathcal{M}}_\tau\left(1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right).
\end{aligned}$$

Subtracting  $\hat{\mathcal{M}}_\tau(n, \hat{q}; t)$  from both sides, dividing by  $\Delta t$ , and take the limit  $\Delta t \rightarrow 0$ , we obtain

$$\begin{aligned}
\frac{\partial \hat{\mathcal{M}}_\tau(1, \hat{q}; t)}{\partial t} &= -\hat{q}(z_I^\tau - \zeta) \frac{\partial \hat{\mathcal{M}}_\tau(1, \hat{q}; t)}{\partial \hat{q}} - (z_X^\tau + \delta_\tau + d_\tau) \hat{\mathcal{M}}_\tau(1, \hat{q}; t) \\
&\quad + 2\delta_\tau \hat{\mathcal{M}}_\tau\left(2, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) + \mu_e m_\tau N_t (1 - \Phi(\hat{q})) \\
&\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \hat{\mathcal{M}}_{\tau'}(1, \hat{q}; t) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \hat{\mathcal{M}}_\tau(1, \hat{q}; t).
\end{aligned}$$

Now  $\mathcal{M}_\tau(1, \hat{q}; t) = \frac{\hat{\mathcal{M}}_\tau(1, \hat{q}; t + \Delta t)}{N_t}$ , thus

$$\begin{aligned}
\frac{\partial \mathcal{M}_\tau(1, \hat{q}; t)}{\partial t} + \eta \mathcal{M}_\tau(1, \hat{q}; t) &= -\hat{q}(z_I^\tau - \zeta) \frac{\partial \mathcal{M}_\tau(1, \hat{q}; t)}{\partial \hat{q}} - (z_X^\tau + \delta_\tau + d_\tau) \mathcal{M}_\tau(1, \hat{q}; t) \\
&\quad + 2\delta_\tau \mathcal{M}_\tau\left(2, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) + \mu_e m_\tau (1 - \Phi(\hat{q})) \\
&\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \mathcal{M}_{\tau'}(1, \hat{q}; t) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \mathcal{M}_\tau(1, \hat{q}; t).
\end{aligned}$$

On a stationary BGP,  $\frac{\partial \mathcal{M}_\tau(1, \hat{q}; t)}{\partial t} = 0$ , so we obtain (28a).

Similarly, for  $n > 1$ ,

$$\begin{aligned}
\hat{\mathcal{M}}_\tau(n, \hat{q}; t + \Delta t) &= \hat{\mathcal{M}}_\tau\left(n, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) - (z_X^\tau + \delta_\tau + d_\tau)\Delta t \hat{\mathcal{M}}_\tau\left(n, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) \\
&\quad + (n+1)\delta_\tau \Delta t \hat{\mathcal{M}}_\tau\left(n+1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) \\
&\quad + (n-1)z_X^\tau \Delta t \hat{\mathcal{M}}_\tau\left(n-1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) \\
&\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \Delta t \hat{\mathcal{M}}_{\tau'}\left(n, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \Delta t \hat{\mathcal{M}}_\tau\left(n, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right).
\end{aligned}$$



Subtracting  $\hat{\mathcal{M}}_\tau(n, \hat{q}; t)$  from both sides, dividing by  $\Delta t$ , and take the limit  $\Delta t \rightarrow 0$ , we obtain

$$\begin{aligned} \frac{\partial \hat{\mathcal{M}}_\tau(n, \hat{q}; t)}{\partial t} &= -\hat{q}(z_I^\tau - \zeta) \frac{\partial \hat{\mathcal{M}}_\tau(n, \hat{q}; t)}{\partial \hat{q}} - (z_X^\tau + \delta_\tau + d_\tau) \hat{\mathcal{M}}_\tau(n, \hat{q}; t) \\ &\quad + (n+1)\delta_\tau \hat{\mathcal{M}}_\tau\left(n+1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) + (n-1)z_X^\tau \hat{\mathcal{M}}_\tau(n-1, \hat{q}; t) \\ &\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \hat{\mathcal{M}}_{\tau'}(n, \hat{q}; t) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \hat{\mathcal{M}}_\tau(n, \hat{q}; t). \end{aligned}$$

Since  $\mathcal{M}_\tau(n, \hat{q}; t) = \frac{\hat{\mathcal{M}}_\tau(n, \hat{q}; t + \Delta t)}{N_t}$ , the last equality implies

$$\begin{aligned} \frac{\partial \mathcal{M}_\tau(n, \hat{q}; t)}{\partial t} + \eta \mathcal{M}_\tau(n, \hat{q}; t) &= -\hat{q}(z_I^\tau - \zeta) \frac{\partial \mathcal{M}_\tau(n, \hat{q}; t)}{\partial \hat{q}} - (z_X^\tau + \delta_\tau + d_\tau) \mathcal{M}_\tau(n, \hat{q}; t) \\ &\quad + (n+1)\delta_\tau \mathcal{M}_\tau\left(n+1, \frac{\hat{q}}{\exp((z_I^\tau - \zeta)\Delta t)}; t\right) + (n-1)z_X^\tau \mathcal{M}_\tau(n-1, \hat{q}; t) \\ &\quad + \sum_{\tau' \neq \tau} \lambda_{\tau'\tau} \mathcal{M}_{\tau'}(n, \hat{q}; t) - \sum_{\tau' \neq \tau} \lambda_{\tau\tau'} \mathcal{M}_\tau(n, \hat{q}; t). \end{aligned}$$

On a stationary BGP,  $\frac{\partial \mathcal{M}_\tau(n, \hat{q}; t)}{\partial t} = 0$ , so we obtain (28b).

The derivation of (27) follows closely Cao and Luo (2017) for wealth distribution with persistent heterogeneous returns.

## C.2 Proofs

# D Robustness exercises

## D.1 One-type model

To check whether two firm types are necessary, we re-estimate the model with one firm type. To do so, we first fix  $\delta = 0.118$  and  $d = 0.004$ .<sup>21</sup> We target  $\eta = 0.01$  and  $\zeta = 0.01$  and the Pareto tail index of the distribution of number of establishment per firm of 1.3060 and the Pareto tail index of the distribution of establishment size of 1.6796, as estimated for 1995.

The Pareto tail index of the distribution of number of establishment per firm is given by (32). Therefore

$$z_X = \frac{\eta + d}{1.3060} + \delta = 0.1287$$

From (18), we have

$$\mu_e = \eta + d + \delta - z_X = 0.0033$$

---

<sup>21</sup>These amount to around 3% quarterly exit rate for establishments and 0.1% quarterly (exogenous) exit rate for firm.

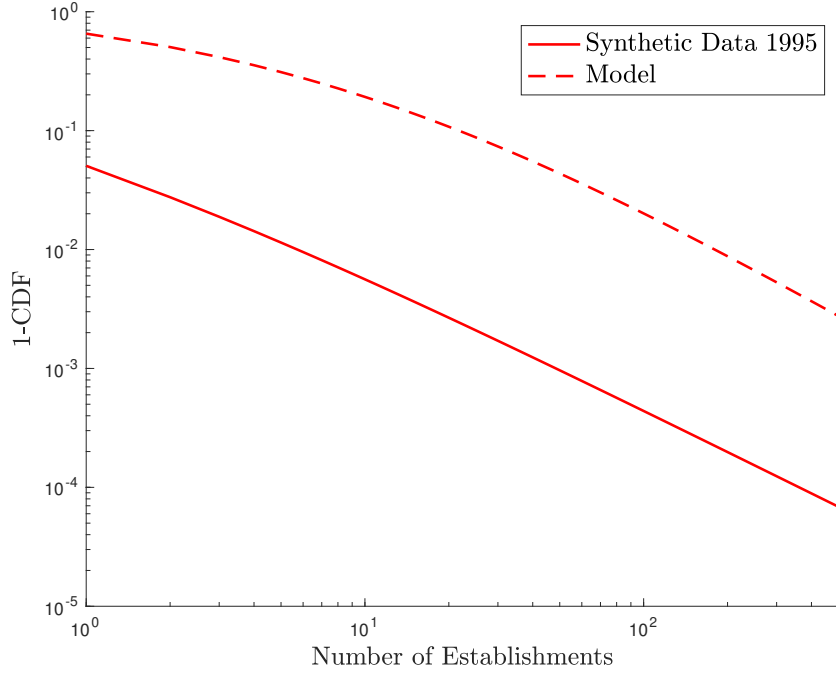


Figure 18: Distribution of number of establishment per firm, Data and Model (One-Type)

The Pareto tail index of the distribution of establishment sizes is given by (30), which implies

$$z_I = \frac{\eta + d + \delta - z_X}{1.6796} + \zeta = 0.0120$$

From (23), we obtain

$$\int \hat{q} d\Phi(\hat{q}) = 1 - \frac{z_I - \zeta}{\mu_e} = 0.4046$$

We assume that  $\Phi$  follows a log-normal distribution with mean  $\varrho$  and variance  $\varsigma^2$ :

$$\Phi \sim \exp(\mathcal{N}(\varrho, \varsigma^2)). \quad (40)$$

Thus

$$\exp\left(\varrho + \frac{\varsigma^2}{2}\right) = \int \hat{q} d\Phi(\hat{q}) = 0.4046$$

Now, the distribution of the number of establishments per firm is given by (31). Figure 18 displays this distribution and shows that this model produces too few firms with one establishments, despite replicating the tail index of the empirical distribution. To better match the data, we therefore need more than one type.

Notably, however, the model generates a distribution of establishment size close to that in the synthetic data at the estimated value of  $\varsigma$ :  $\hat{\varsigma} = 1.2606$ . (and  $\hat{\varrho} = -1.6993$ ). The model implied

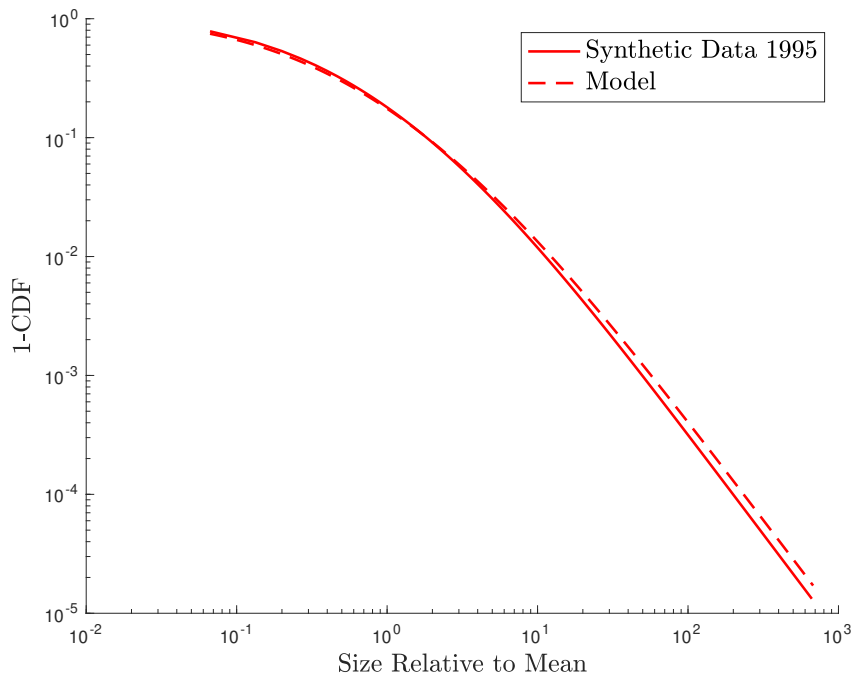


Figure 19: Distribution of number of employee per establishment, Data and Model (One-Type)

distribution as given by (29) and the empirical distribution are shown in Figure 19.