

# The Reallocation Myth

Chang-Tai Hsieh

Peter J. Klenow \*

University of Chicago and NBER

Stanford University and NBER

March 31, 2018

## Abstract

There is a widely held view that much of growth in the U.S. can be attributed to reallocation from low to high productivity firms, including from exiting firms to entrants. Declining dynamism — falling rates of reallocation and entry/exit in the U.S. — have therefore been tied to the lackluster growth since 2005. We challenge this view. Gaps in the return to resources do not appear to have narrowed, suggesting that allocative efficiency has not improved in the U.S. in recent decades. Reallocation can also matter if it is a byproduct of innovation. However, we present evidence that most innovation comes from existing firms improving their own products rather than from entrants or fast-growing firms displacing incumbent firms.

---

\*Prepared for the 2017 Federal Reserve Bank of Kansas City Economic Policy Symposium on *Fostering a Dynamic Global Economy*. We are grateful to Mathias Jimenez, Matteo Leombroni, and Jihoon Sung for excellent research assistance. Steve Davis, Gita Gopinath, and John Haltiwanger provided constructive criticism. Hsieh acknowledges support from the Polsky Center at the University of Chicago's Booth School of Business, and Klenow from the Stanford Institute for Economic Policy Research. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

## 1 Introduction

A widespread view holds that reallocation from less to more productive firms (including from exiting firms to entrants) is central to growth. Baily et al. (1992) and Foster et al. (2001) are classic references. A corollary is that declining rates of firm entry and job reallocation have contributed substantially to the lackluster growth seen in the U.S. for at least a decade (Decker et al., 2014).

It is not clear, however, that resources should continually be reallocated from less productive to more productive firms. If firms produce heterogeneous products, then it is not optimal to pile all inputs in the economy onto the highest productivity firm. In many environments, it is optimal to stop reallocating when the marginal return to resources is equalized across firms. Reallocation improves allocative efficiency only when resources move from firms where the *marginal* revenue product of resources is low to firms where the *marginal* revenue product is high. See Hsieh and Klenow (2009) for an example environment in which this is necessary condition for allocative efficiency.

After clarifying this point, we present evidence that allocative efficiency has not improved in U.S. manufacturing in recent decades. Less complete evidence outside manufacturing likewise suggests no positive contribution of allocative efficiency to U.S. growth. How is this evidence reconciled with the conventional growth accounting? Our answer is that reallocation which services to lower gaps in marginal revenue products is offset by shocks that increase such gaps. The net effect is that allocative efficiency has not improved or has even worsened, so that its apparent contribution to growth is illusory.

A variant of the reallocation-growth nexus is that innovation requires reallocation of inputs. This is true, for example, in Schumpeter (1939)'s vision of creative destruction as the driving force behind growth. Here, reallocation can be important with no change in allocative efficiency. Indeed, reallocation can contribute to growth without any gaps in marginal revenue products at all. Instead, what is essential is that inputs reallocate from firms that fail to innovate

to firms that do innovate.

How much growth comes from creative destruction? To provide evidence on this question, in joint work with Daniel Garcia-Macia we exploit data on job flows across firms (Garcia-Macia et al., 2016). The basic idea is that creative destruction should show up as big declines (and, in the extreme, exit) among some incumbents and major growth among those firms doing the destroying. That is, creative destruction should show up in the “tails” of the distribution of employment changes across firms.

In contrast, when incumbents improve their own products they replace themselves, entailing much smaller changes in firm employment. Think of a retailer upgrading an outlet as opposed to opening a new store and driving out a competitor. Such incremental innovations within firms should populate the middle of the job growth distribution. Since most employment changes across private firms in the U.S. are modest, even over five year periods, this tells us that most innovation takes the form of incumbents improving their own products.

To summarize, we reach two main conclusions about U.S. growth in recent decades. First, we see no improvement in allocative efficiency. There has been reallocation from low to high marginal revenue product firms, but evidently new shocks have prevented gaps in marginal revenue products from diminishing over time. Thus U.S. growth is not driven by reallocation that enhances allocative efficiency. Second, reallocation from firms that did not successfully innovate to firms that did accounts for a modest fraction (about a quarter) of U.S. aggregate productivity growth. So the second notion of reallocation, from less innovative to more innovative firms, cannot explain most aggregate growth according to our evidence.

The rest of the paper proceeds as follows. Section 2 describes the traditional growth decomposition and why it is at odds with a benchmark model of heterogeneous firms. Section 3 presents a richer model with distortions and evidence that allocative efficiency has not contributed positively to U.S. growth in recent decades. Section 4 describes our alternative growth decomposition, which ar-

rives at a dominant role for incumbent innovation on their own products rather than entrants and fast-growing firms (so-called “gazelles”). Section 5 concludes.

## 2 Reallocation from low to high productivity firms

There is a widespread view that a significant part of aggregate growth comes from the reallocation of resources from low to high productivity firms. A widely used accounting exercise, introduced by Baily et al. (1992), attempts to measure this mechanism. Specifically, in a model where labor is the only factor of production, Baily et al. (1992) decompose growth in aggregate output per worker  $y$  into:

$$\Delta \ln y_t = \sum_{i \in C} s_{i,t-1} \Delta \ln y_{i,t} + \sum_{i \in C} \Delta s_{i,t} \ln y_{i,t} + \sum_{i \in E} s_{i,t} \ln y_{i,t} - \sum_{i \in X} s_{i,t-1} \ln y_{i,t-1} \quad (1)$$

where  $C$ ,  $E$ , and  $X$  denote the set of incumbent firms (firms who produce in both  $t - 1$  and  $t$ ), entering firms in year  $t$ , and exiting firms in year  $t - 1$ , respectively.  $y_i$  is revenue per worker of firm  $i$ , and  $s_i$  is the revenue share of firm  $i$ .<sup>1</sup> The first term in equation (1) is often interpreted as the effect of “within-firm” productivity growth and the last three terms are often interpreted as the contribution of reallocation from low to high labor productivity firms.

The intuition behind equation (1) is that aggregate output per worker is a geometric-weighted average of revenue per worker of individual firms (deflated by an aggregate price index). Therefore, aggregate output increases when the market share of high labor productivity firms increases (the second term), when entrants have high labor productivity (the third term), and when exiting firms have low labor productivity (the fourth term). Note that “labor productivity” here is revenue per worker divided by an aggregate deflator.

In U.S. manufacturing, high labor productivity firms typically gain employ-

---

<sup>1</sup>In practice, the decomposition is often done at the industry level, and sometimes with Total Factor Productivity (TFP) rather than labor productivity, or at the plant level rather than the firm level.

ment share, and entrants have higher labor productivity than exiting firms on average. Baily et al. (1992) estimate that the reallocation of resources towards high productivity incumbent firms — the second term in (1) — boosted growth by an average of 0.5 percentage points per year between 1972 and 1987 in five manufacturing industries, equal to about one-third of their overall growth. They found no average contribution from net entry. Foster et al. (2001) provide a range of estimates for the contribution of reallocation to productivity growth in all of U.S. manufacturing from 1977–1992, with the second term in (1) averaging 18% of growth. They found a further 22% contribution from net entry, and 60% from within-firm productivity growth.

It is difficult to interpret accounting decompositions such as (1), however, in the absence of an equilibrium model of heterogeneous firms. Consider, for example, how statistics based on (1) would map into a model of heterogeneous firms who are monopolistic competitors facing constant elasticities of substitution (CES demand). Here, aggregate output  $Y$  is a CES combination of the output of individual firms  $Y_i$ :

$$Y = \left( \sum_{i=1}^N Y_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \quad (2)$$

Suppose further that firm output is given by  $Y_i = A_i \cdot L_i$  where  $A_i$  denotes firm-specific total factor productivity (TFP) and  $L_i$  denotes labor input. Assuming firms maximize profits, the employment share  $l_i$  and labor productivity  $y_i$  of a firm are given by

$$l_i \equiv \frac{L_i}{L} = \left( \frac{\sigma-1}{\sigma} \cdot \frac{A_i}{w} \right)^{\sigma-1} \quad (3)$$

$$y_i \equiv \frac{P_i Y_i}{L_i} = \frac{\sigma-1}{\sigma} \cdot w \quad (4)$$

where  $w$  is the wage.<sup>2</sup> Equations (3) and (4) illustrate two points. First, labor productivity (revenue per worker) is the same across all firms and is not a measure

---

<sup>2</sup>We normalize the aggregate price index to one so  $w$  is the real wage.

of firm TFP ( $A_i$ ).<sup>3</sup> Second, differences in firm TFP ( $A_i$ ) show up as differences in employment. High TFP firms expand until the marginal revenue product of labor is equal to the common wage. With CES demand and monopolistic competition, the marginal revenue product is proportional to the average revenue product (labor productivity).

After imposing the condition that aggregate labor demand is equal to aggregate labor supply, aggregate labor productivity is given by

$$y = \left( \sum_{i=1}^N A_i^{\sigma-1} \right)^{\frac{1}{\sigma-1}}.$$

The growth rate of aggregate labor productivity is then given by:<sup>4</sup>

$$\Delta \ln y_t = \sum_{i \in C} \tilde{l}_{it} \cdot \Delta \ln A_{i,t} + \frac{1}{\sigma-1} \ln \sum_{i \in E} l_{i,t} - \frac{1}{\sigma-1} \ln \sum_{i \in X} l_{i,t-1}. \quad (5)$$

Compare the sources of growth implied by the equilibrium model in equation (5) with the accounting decomposition by Baily et al. (1992) in (1).<sup>5</sup> First, there are no gains from reallocating resources from low TFP to high TFP firms. This is because the marginal revenue product of labor is the same in all firms.<sup>6</sup> High TFP firms use more labor, but as long as the marginal return to labor is the same across all firms, aggregate output *falls* if more resources are allocated to high TFP firms (say if high TFP firms were to be subsidized and low TFP firms

---

<sup>3</sup>Even though firm output is divided by the aggregate price level, there is no variation in deflated revenue per worker here. A firm-specific price is needed to uncover firm-specific process efficiency. This is done by Foster et al. (2008) and others. A firm-specific unit price will not, however, reveal firm differences in quality and variety. See Hottman et al. (2016) for evidence that firm size differences in consumer goods manufacturing entirely reflects quality and variety differences.

<sup>4</sup>The weight  $\tilde{l}_{it} \equiv \frac{\Delta l_{i,t}}{\Delta \ln l_{i,t}} / \sum_{j=1}^N \frac{\Delta l_{j,t}}{\Delta \ln l_{j,t}}$  in equation (5) is the Sato-Vartia average of employment shares in  $t$  and  $t-1$ . See Sato (1976) and Vartia (1976).

<sup>5</sup>The revenue share of the firm in this model is proportional to the employment share. Therefore, we could have just as easily used the revenue share, as in the Baily et al. (1992) decomposition (1), for the weights in (5).

<sup>6</sup>Firms do charge markups here, but every firm faces the same elasticity of demand and charges the same markup. Thus there is no misallocation of labor across firms.

were to be taxed). Second, the contribution of entrants to growth is measured by their *employment* share, not by their labor productivity.<sup>7</sup> Third, exit *harms* growth, where the sufficient statistic for the loss is the employment share of exiting firms.<sup>8</sup>

Empirically labor productivity differs across firms even within narrow industries (Syverson, 2011). And, as mentioned, there is abundant evidence in the U.S. that inputs tend to move toward high labor productivity firms. The equilibrium model behind (5) does not capture this fact. To explain this empirical regularity, we have to relax some of the assumptions that imply an identical return to labor across all firms. For example, suppose that firms have to choose employment one period in advance, before observing their level of TFP ( $A_i$ ). If TFP follows a random walk then labor productivity becomes:<sup>9</sup>

$$y_{i,t} = \frac{\sigma - 1}{\sigma} \cdot w_t \cdot \left( \frac{A_{i,t}}{A_{i,t-1}} \right)^{\frac{\sigma-1}{\sigma}} . \quad (6)$$

Labor productivity (revenue per worker) now differs across firms because of lags in adjusting employment to productivity shocks. Note, however, that labor productivity still does not reflect differences in TFP, but rather the *change* in TFP.

With this modification, the equilibrium model can now explain, at least qualitatively, the empirical patterns in decompositions based on equation (1). First, labor productivity is high among firms with positive TFP shocks and low in firms with negative TFP shocks. In the absence of new TFP shocks, high labor productivity firms gain employment and low productivity firms lose employment. Reallocation thus increases aggregate output because it narrows the gap in the return to labor across firms. The gain is not due to increasing the market share of high TFP firms *per se*. Reallocation is driven by “within” firm productivity shocks. In the absence of within-firm productivity changes, there would be no

---

<sup>7</sup>If we added capital to the model, then the entrant overall input share would matter.

<sup>8</sup>Even low TFP varieties are valuable given love of variety. Of course, exit could be valuable in a richer model with overhead costs or creative destruction by a close substitute.

<sup>9</sup>More generally, the denominator would be  $E(A_{i,t}|A_{i,t-1})$  instead of  $A_{i,t-1}$ .

reallocation.<sup>10</sup>

Second, if it also takes time for entrants to reach their steady state size, then labor productivity of entrants is high as long as they are below their frictionless size.<sup>11</sup> The ultimate contribution of entrants to aggregate growth is still captured by their employment share, however, not by their labor productivity. High labor productivity implies that entrants are below the frictionless size implied by their TFP. If entrants could adjust employment faster, their labor productivity would be lower, their employment share would be higher, and their contribution to growth would be even larger.

Bartelsman et al. (2013) add overhead labor to this environment, generating an additional source of dispersion in revenue per worker. But overhead costs that are identical across all firms cannot be very large because the smallest firm has only one worker in most four-digit industries. Moreover, overhead costs generate variation in labor productivity only as a result of variation in firm size (employment). Hsieh and Klenow (2009) report that size accounts for very little of the variation in revenue per unit of inputs across firms in the U.S., China and India. Age likewise explains little of the variation in labor productivity.

Decker et al. (2017b) simulate a model with a very high correlation (around 0.9) between TFP and labor productivity across plants. But in the U.S. data these variables are far from synonymous. In Hsieh and Klenow (2014), we estimate a correlation of only 0.16 between revenue relative to inputs and our underlying measure of TFP across plants. We find an elasticity of revenue/inputs with respect to TFP of only 0.09. We find much higher elasticities in India (0.50) and Mexico (0.66).

In sum, the gains from reallocating inputs from low to high labor productivity firms do not come from increasing the market share of high TFP firms *per*

---

<sup>10</sup>Adjustment costs affect the level of aggregate productivity, not its growth rate, in this model. Asker et al. (2014) present evidence that countries with big firm-level productivity shocks exhibit more dispersion in revenue per unit of capital.

<sup>11</sup>See Gourio and Rudanko (2014) and Foster et al. (2016) for models and evidence on customer base accumulation as a force for reallocation from established to young firms.

se in our environment. Instead, reallocation contributes to aggregate output because it narrows the gap in the return to labor across firms. Such dispersion in labor productivity can be driven by the creation of new products (entry) and productivity shocks among incumbent firms combined with frictions in reallocating inputs. But gaps in labor productivity dissipate as inputs are reallocated in response to these shocks. In this environment, reallocation is an endogenous response to productivity, not an independent force behind aggregate growth. If productivity shocks are ongoing so that dispersion in labor productivity among survivor firms is stable, then in our environment reallocation among firms contributes to the level of aggregate productivity but not its growth rate.

### 3 Allocative efficiency

Reallocation can be an independent contributor to aggregate growth, at least for awhile, if differences in the return to labor are *shrinking* over time. This might happen, for example, if differences in labor productivity are driven by firm-specific taxes or subsidies that become less dispersed over time. The reallocation in response to such shrinking tax rate dispersion increases aggregate output by narrowing the gap in the return to resources across firms.<sup>12</sup>

In the environment sketched above, aggregate labor productivity  $y$  is given by:

$$y = \underbrace{\left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{A_i}{\bar{A}} \right)^{\sigma-1} \left( \frac{\tau_i}{\tau} \right)^{1-\sigma} \right]^{\frac{1}{\sigma-1}}}_{AE=\text{Allocative Efficiency}} \times \underbrace{\left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{A_i}{\bar{A}} \right)^{\sigma-1} \right]^{\frac{1}{\sigma-1}}}_{PD=\text{Productivity Dispersion}} \times \underbrace{N^{\frac{1}{\sigma-1}}}_{\text{Variety}} \times \bar{A}$$

<sup>12</sup>Differences in labor productivity could also arise from markup dispersion as in Peters (2016), Haltiwanger et al. (2016), and Autor et al. (2017). Changing dispersion in markups could generate trends in allocative efficiency that affect aggregate growth, at least temporarily. See Hopenhayn (2014) and Restuccia and Rogerson (2017) for recent surveys of the causes and consequences of productivity dispersion and misallocation.

where

$$\begin{aligned}\tilde{A} &\equiv \left[ \frac{1}{N} \sum_{i=1}^N A_i^{\sigma-1} \right]^{\frac{1}{\sigma-1}} && \text{(Generalized mean productivity)} \\ \bar{A} &\equiv \prod_{i=1}^N A_i^{\frac{1}{N}} && \text{(Geometric mean productivity)} \\ \tau &\equiv \left( \sum_{i=1}^N \frac{P_i Y_i / \tau_i}{PY} \right)^{-1} && \text{(Generalized mean gross tax rate)}\end{aligned}$$

Here  $\tau_i$  is a gross revenue tax rate (if  $> 1$ ) or subsidy rate (if  $< 1$ ). The allocative efficiency term captures the impact on aggregate productivity of dispersion in the return to labor across firms due to dispersion in  $\tau$ .

Ignoring entry and exit, aggregate labor productivity growth can be decomposed as

$$\Delta \ln y_t = \underbrace{\Delta \ln \bar{A}_t}_{\text{Average Productivity Contribution}} + \underbrace{\Delta \ln PD_t}_{\text{Productive Dispersion Contribution}} + \underbrace{\Delta \ln AE_t}_{\text{Allocative Efficiency Contribution}} \quad (7)$$

The first two terms are conceptually akin to the within-productivity term in the conventional decomposition. The decomposition here clarifies that there are two distinct contributions — an increase in the unweighted mean level of firm TFP, and a contribution from any increase in the dispersion of firm level TFP's. Dispersion in TFP is good because  $\sigma > 1$  and more inputs can be allotted to the higher TFP firms (up to a point). The last term captures how aggregate productivity rises in response to falling dispersion in  $\tau_i$ .<sup>13</sup> There is no direct role for reallocation of inputs from low to high TFP firms.

Note that when there are firm specific taxes or subsidies, labor productivity of a firm is given by:

$$y_i = \frac{\sigma - 1}{\sigma} \cdot w \cdot \tau_i.$$

<sup>13</sup>When  $A_i$  and  $\tau_i$  are jointly log-normally distributed, the allocative efficiency contribution is simply  $-(\sigma - 1)/2 \cdot \Delta \text{variance}(\ln \tau_i)$ . See Hsieh and Klenow (2009).

As before, labor productivity (revenue per worker) does not measure firm TFP ( $A_i$ ). High labor productivity indicates that the firm faces a high effective tax rate. The dispersion in labor productivity (revenue relative to inputs) thus reflects the dispersion in the tax rate. Allocative efficiency improves when the dispersion of labor productivity falls.

A widely used method by Olley and Pakes (1996) measures allocative efficiency by the covariance between firm inputs and firm labor productivity (revenue per worker).<sup>14</sup> A higher covariance is interpreted as indicating higher allocative efficiency. In our model, allocative efficiency is at the maximum when labor productivity is the same in all firms and the covariance between labor productivity and size is *zero*. Relative to the efficient benchmark, a higher covariance of size and labor productivity would actually worsen allocative efficiency. More generally, it is not the covariance of size and labor productivity that matters for allocative efficiency, but rather the dispersion of labor productivity.

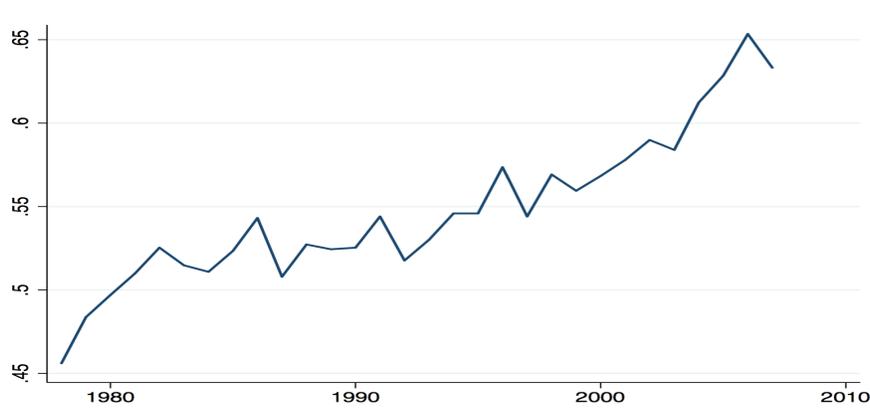
In Hsieh and Klenow (2009), we measure the change in allocative efficiency in China by the change in the weighted dispersion in revenue per unit of inputs. This evidence suggests that rising allocative efficiency in China accounted for about one-third of overall manufacturing productivity growth from 1998–2007. We trace about two-fifths of better allocative efficiency to the shrinking role of state-owned firms with low revenue per unit of inputs. Bai et al. (2016) show that the improvement in allocative efficiency was partially undone after 2010 when Chinese local governments were able to use off-balance sheet investment companies (known as “local financing vehicles”) to fund favored private firms.

In U.S. manufacturing, we see no evidence that allocative efficiency has improved. Figure 1 plots the standard deviation of revenue per unit inputs from the Annual Survey of Manufacturing. Consistent with the evidence in Kehrig (2015), dispersion has risen in recent decades. Bils et al. (2017) calculate that the implied fall in allocative efficiency would be big enough to cut manufacturing

---

<sup>14</sup>See Decker et al. (2017a) for a recent example of growth accounting using the Olley-Pakes decomposition.

Figure 1: Dispersion in revenue/inputs in the U.S.

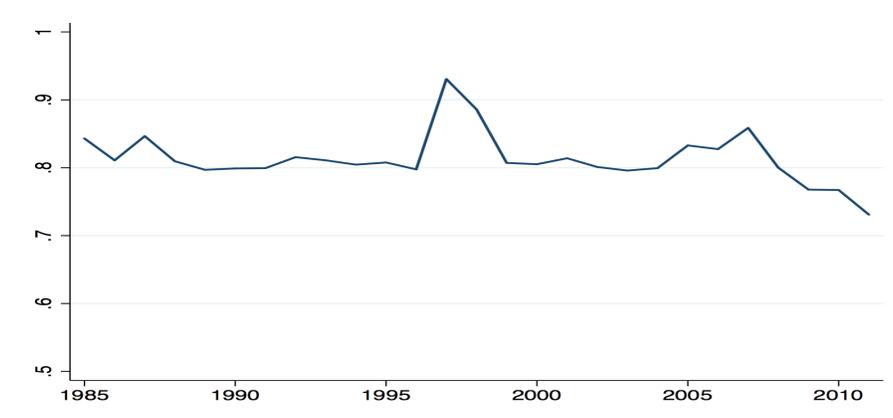


Source: The U.S. Annual Survey of Manufacturers, 1978–2007. The plot shows the average within-industry standard deviation of log (revenue/inputs) across plants. The weights are industry value added shares.

TFP in half between 1978 and 2007. They raise the possibility, however, that this trend is benignly due to rising measurement error or model misspecification. When they correct for such errors, they conclude that allocative efficiency was largely unchanged. As for the U.S., they estimate that allocative efficiency changed little in India from 1985–2011. See Figure 2. Gopinath et al. (2017), however, find that falling allocative efficiency, in particular from rising dispersion in revenue/capital ratios, lowered growth about 1% per year in Spanish manufacturing from 1999 to 2012.

Evidence for the broader U.S. economy points to rising dispersion in labor productivity. Using data from the Economic Censuses from 1977 to 2007, Barth et al. (2016) document that dispersion in revenue per worker rose twice as fast within six sectors outside manufacturing as within manufacturing. Decker et al. (2017b) report similar patterns in administrative data, suggesting the adverse trends are real rather than purely measurement error. Hsieh and Moretti (2017) suggest that the rising dispersion in revenue per worker is partially driven by increasingly stringent housing constraints in coastal U.S. cities.

Figure 2: Dispersion in revenue/inputs in India



Source: India's Annual Survey of Industries, 1985–2011. The plot shows the average within-industry standard deviation of  $\log(\text{revenue}/\text{inputs})$  across plants. The weights are industry value added shares.

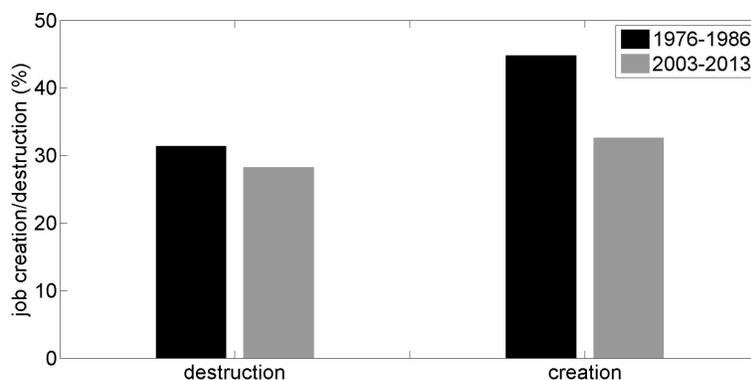
To recap, traditional growth decompositions indicate gains from reallocation from low labor productivity to high productivity firms. But if dispersion in revenue productivity is stable or rising, then changing allocative efficiency would not appear to be positive source of growth at all. This is the case in U.S. manufacturing in recent decades.

## 4 Innovation and Reallocation

Figure 3, reproduced from Garcia-Macia et al. (2016), presents the aggregate job creation and destruction rates over a five year period among all private non-farm businesses in the U.S. The magnitude of the reallocation of labor shown in Figure 3 is enormous, and most of this reallocation is not associated with shifts from low to high labor productivity firms. This is consistent with the classic findings of Davis et al. (1998) for manufacturing.

Suppose that innovating firms replace products made by other firms as in Schumpeter (1939)'s growth through creative destruction. If this is the dom-

Figure 3: Job Creation and Destruction Rates in U.S.



Source: U.S. Census Longitudinal Business Database (LBD) on firms in the nonfarm business sector. The job creation (destruction) rate is the sum of employment changes at firms with rising (falling) employment divided by aggregate employment in the initial period. This includes entering and exiting firms. The figure shows the average 5-year changes for 1976–1981 and 1981–1986, and for 2003–2008 and 2008–2013.

inant source of growth, then the large flows of labor depicted in Figure 3 are a necessary byproduct of innovation – even if they do not involve reallocation from low to high revenue productivity firms. In this view, the firms whose products have been creatively destroyed are behind the job destruction rates displayed in Figure 3. They need not have been low revenue productivity firms.

Innovation would come to a halt if reallocation was not possible, the creative destruction argument goes. Acemoglu and Robinson (2012), for example, provide historical accounts of countries who stopped growing when creative destruction was blocked. In the case of the U.S., Decker et al. (2014) point to the decline in the job creation rate since the late 1970s, which can be seen in Figure 3, as evidence of a decline in U.S. innovation.

In Garcia-Macia et al. (2016), we estimate the importance of creative destruction for overall growth. We use the equilibrium model of heterogeneous firms in section 2, except that we interpret  $i$  in equation (2) as an index of a *product* and not a firm. This distinction is crucial because, when firms produce

multiple products, we can no longer directly map reallocation of employment at the firm level into innovation at the product level. It is innovation at the product level which drives aggregate growth in our model. For simplicity, assume the aggregate number of products is fixed.<sup>15</sup> The growth rate is then a weighted average of the TFP growth rates of each product  $i$ :

$$\Delta \ln y_t = \sum_{i=1}^M \tilde{l}_{it} \Delta \ln A_{i,t} \quad (8)$$

where  $l_i$  is the share of labor used in making product  $i$  and  $M$  is the fixed number of total products (which is greater than the number of firms  $N$ ).<sup>16</sup>

The firm with the highest TFP for each product takes over the market for that product. The number of products owned by a firm is thus the number of products for which it has the highest TFP. The employment share of firm  $f$  is then the sum of the employment share of the products it owns:

$$l_f = \sum_{i \in f} l_i = \left( \frac{\sigma - 1}{\sigma} \cdot \frac{1}{w} \right)^{\sigma - 1} M_f \bar{A}_f^{\sigma - 1} \quad (9)$$

where  $M_f$  is the number of products firm  $f$  owns and  $\bar{A}_f \equiv \left( \frac{\sum_{i \in f} A_i^{\sigma - 1}}{M_f} \right)^{\frac{1}{\sigma - 1}}$  is the generalized mean TFP of its products.

So far this structure mimics Klette and Kortum (2004), wherein a firm is a portfolio of products and all growth comes from creative destruction. But a key question for job reallocation is who is improving TFP at the product-firm level. TFP growth at the product level in equation (8) can come from innovation by the firm that already owns the product (“own innovation”) or from another firm (“creative destruction”). Creative destruction, in turn, can be carried out by an-

<sup>15</sup>In Garcia-Macia et al. (2016) we allow for the creation of new products, but we estimate this is not a major source of growth (< 10% of all growth).

<sup>16</sup>Recall that  $\tilde{l}_i$  is the Sato-Vartia average of the employment share of product  $i$  in the two time periods ( $t$  and  $t - 1$ ), where the employment share  $l_i$  is given by equation (3). Note that equation (8) is the same as equation (5) except with the last two terms set to zero because we assume no entry and exit of products.

other incumbent or by an entrant. Although an instance of own innovation and creative destruction could have the exact same effect on aggregate productivity, their effect on reallocation can be very different.

Consider the effect of innovation on reallocation in a toy economy with four products with identical TFP (initially) and two firms (Incumbent 1 and Incumbent 2), each with two products. Suppose further that the quality of one of the products owned by Incumbent 2 improves by 8% so that aggregate output (and the average wage) rises by 2%.<sup>17</sup> Table 1 shows that this innovation has a very different effect on employment reallocation depending on who carries out the innovation. The first row assumes the innovator is Incumbent 1, who did not previously own the product (creative destruction). The employment share of the innovating firm (Incumbent 1) increases by 26 percentage points. The employment share of the firm whose product was innovated upon (Incumbent 2) drops by the same amount so the economy's job creation and destruction rates are both 26 percent.<sup>18</sup> The reallocation is large because the employment gain of the innovating firm mostly comes from stealing the business of the other firm. In equation (9), the number of products  $M$  of Incumbent 1 increases by 50% (from 2 to 3) while the number of products of Incumbent 2 drops by 50% (from 2 to 1). The business stealing effect accounts for a 25 percentage point shift in the employment share. The total shift in the employment share is 26 percentage points — the additional percentage point comes from the fact that the product was improved, not just stolen.

The second row considers the employment shifts when the innovator is a brand new firm (an “entrant”). This is still creative destruction, except that the innovating firm previously did not exist. The effect on the employment share of Incumbent 2 is the same, a 26 percentage point drop. The difference is that employment of Incumbent 1 also drops by a small amount (by 2 percentage

<sup>17</sup>From (8), the aggregate growth rate is  $.25 \cdot .08 = .02$ , or 2%.

<sup>18</sup>Using (9), the ratio of employment of Incumbent 1 to Incumbent 2 is  $\frac{M_1}{M_2} \cdot \left(\frac{A_1}{A_2}\right)^{\sigma-1} = \frac{3}{1} \cdot \frac{1.0555}{1} = 3.155$ . (We assume  $\sigma = 3$ ). The new employment share of the innovator is  $\frac{3.155}{3.155+1} = 0.76$ .

Table 1: Effect of Innovation on Reallocation in a Toy Economy

	Δ Employment Share		
	Entrant	Incumbent 1	Incumbent 2
Creative Destruction by Incumbent 1	–	+26%	-26%
Creative Destruction by Entrant	+28%	-2%	-26%
Own Innovation by Incumbent 2	–	-2%	+2%

Note: Before innovation, there are two firms (Incumbent 1 and Incumbent 2) each with two products, and all four products are of equal importance. The innovation is an 8% improvement in a single product, boosting aggregate output by 2%. The entries illustrate how job reallocation differs depending on who carries out the 8% improvement.

points) due to the general equilibrium effect of a rising wage. Job creation is driven by the entrant, whose employment share is 28 percent. So the aggregate job creation (and destruction) rate when an entrant innovates is 28 percent.

Now consider the effect of the same innovation, except that innovation is undertaken by the firm that already owns the product (incumbent 2). The effect on aggregate output is the same, but the effect on reallocation is much smaller. The third row in Table 1 shows that the employment share of the innovating firm rises by only 2 percentage points; its average quality  $\bar{A}$  rises by 4.1 percent and its number of products  $M$  does not change.<sup>19</sup> Employment falls by 2 percentage points in the firm that failed to innovate — in equation (9) this is the effect of the rising wage on employment. The aggregate job creation (and destruction) rate in this case is only 2 percent.

In short, the magnitude of reallocation by itself does not tell us how much innovation and growth has occurred. It is more informative about the *relative* contribution of creative destruction vs. own innovation to growth. Garcia-Macia et al. (2016) exploit these divergent implications of creative destruction

<sup>19</sup>From (9), the ratio of employment at Incumbent 2 (the innovating firm) to Incumbent 1 is  $\frac{M_2}{M_1} \cdot \left(\frac{\bar{A}_2}{\bar{A}_1}\right)^{\sigma-1} = \frac{2}{2} \cdot \left(\frac{1.041}{1}\right)^2 = 1.083$ . The new employment share of incumbent 2 is  $\frac{1.083}{1.083+1} = 0.52$ .

vs. own innovation for the distribution of employment reallocation across firms. Do most firms undergo large increases in employment — consistent with growth driven by creative destruction — or do they experience only “small” changes in employment, consistent with growth through own innovation? We find that the vast majority of expanding firms grow modestly, which suggests that innovation largely takes the form of firms improving the quality of their own products.<sup>20</sup>

Table 2 presents our estimates of the contribution of job reallocation to growth over the 2003–2013 period. The first column presents the contribution of entrants, defined as firms established in the last five years. The first row gives their contribution to reallocation and the second their contribution to aggregate growth. We infer that entrants account for 50% of the 30% aggregate job creation rate seen in Figure 3 over the 2003–2013 period (the employment share of entrants is about 15%). Since innovation by entrants largely takes the form of creative destruction rather than new products, the employment share of entrants is mostly due to business stealing from incumbent firms as opposed to the quality improvements they make. As a result, the contribution of entrants to aggregate growth is smaller than their share of job creation. We estimate that entrants accounted for 13% of aggregate growth over the 2003–2013 period.

The second column of Table 2 presents our estimates of the contribution of fast-growing firms, defined as firms that grow by more than 20% per year over a five year period. Such firms are the “gazelles” that many view as having a disproportionate contribution to growth.<sup>21</sup> Our interpretation however, is

---

<sup>20</sup>One might worry that demand shocks, rather than innovation, are responsible for the modest growth rates of many firms. But note that we are averaging growth across five-year periods, so that temporary demand shocks should largely wash out. The same argument goes for adjustment costs, whose effects should be mitigated by taking five year averages. Another concern might be that modest growth rates reflect secular shifts due to nonhomothetic preferences, such as from agriculture and manufacturing to services. But secular shifts seem to occur almost entirely on the extensive margin (new firms) rather than within-firms — see Bollard et al. (2016). And the vast majority of job reallocation occurs within narrow industries, as emphasized by Davis et al. (1998). To the extent that demand shocks occur over five-year periods within industries, the open question is whether they are thick-tailed (mimicking creative destruction) or modest (thereby mimicking own innovation).

<sup>21</sup>This term seems to have been coined by Birch (1981). See also Acs and Mueller (2008). We based the threshold of 20% per year growth on Ahmad and Gonnard (2007).

Table 2: Contribution to Aggregate Growth and Job Creation, 2003-2013

	Entrants	Incumbents by Job Growth Rate		
		> 20%	0% to 20%	< 0%
Share of Job Creation	50.3%	13.2%	36.5%	–
Share of Aggregate Growth	13.3%	4.3%	64.6%	17.8%

Sources: The first row uses data from the U.S. LBD on firms in the nonfarm business sector. The second row is based on the indirect inference in Garcia-Macia et al. (2016).

that the rapid growth of gazelles is largely due to creative destruction. This is because the right tail of job creation is paired with an (almost equally thick) left tail of job destruction. Because only a modest fraction of job creation (13%) is due to gazelles, our calculation is that their contribution to aggregate growth is quite small. In row 2 we attribute only about 4% of aggregate productivity growth from 2003–2013 to innovation by such rapidly growing firms.<sup>22</sup>

As displayed in the third column of Table 2, we estimate that the main driver of U.S. growth is comparatively slow-growing firms, i.e., those growing less than 20% per year over a five year period. Such firms account for 36% of total job creation, but are responsible for almost two-thirds of aggregate productivity growth. Again, this is because slow growing firms innovate by improving their own products, which has only a minor effect on job growth. The Apples, Intels, and GMs of the world largely replace themselves when they introduce new versions of their existing products.

Finally, the last column of 2 presents the growth contribution of firms that *decline* in size. In our framework, firms shrink when their products are taken over by other firms. But even when this occurs, it might still be the case that they have innovated on at least some of their products and/or creatively destroyed some products. Put differently, although many firms that shrink do not

<sup>22</sup>The contrast is even starker if, as is sometimes done, one argues that gazelles account for more than 100% of *net* employment growth in the economy.

**Table 3:** Share of Growth due to Creative Destruction

	Creative Destruction by	
	Entrants	Incumbents
1976–1986	19.1%	8.2%
2003–2013	12.5%	6.4%

Source: Indirect inference in Garcia-Macia et al. (2016).

innovate, some of these firms may shrink simply because they do not innovate *enough* relative to other firms. Although they do not contribute as much as other firms, their innovation still matters for growth. Table 2 shows that firms that shrink are responsible for almost 18% of aggregate growth over this period, in our estimation. Although shrinking firms are rarely glorified, we estimate that their innovation contributes more to growth than innovation by entrants, and more than four times more than innovation by the fast growing gazelles.

To be clear, we do not mean to imply that reallocation does not matter for growth. Examples abound such as Apple and Samsung smartphones creatively destroying Nokia and Blackberry, Wal-Mart creatively destroying mom-and-pop shops, and Amazon creatively destroying brick-and-mortar retailers. Aggregate growth would surely be lower if creative destruction were blocked. It would also be lower if there were no reallocation in response to own innovation. Our claim instead is that, although most of the job growth (and destruction) in the data is due to innovation by entrants and fast growing “gazelles”, their contribution to growth is notably smaller than their contribution to job reallocation. The main driver of U.S. innovation appears to be comparatively slow-growing firms that have a much smaller effect on job creation.

In addition to average growth, there is the question of whether the declining dynamism documented by Decker et al. (2014) has contributed to the slow U.S. growth in the past decade. Table 3 presents our estimates of the contribution

of creative destruction to growth. The estimated contribution of growth from entrants and creative destruction fell along with the rates of firm entry and job creation. The contribution of entrants fell by six percentage points and that from creative destruction by eight percentage points. Since entrants and creative destruction were modest contributors to growth, however, this declining dynamism shaved off only 10-15 basis points from the annual growth rate, *ceteris paribus*. This is not trivial, but would not explain much of the 180 basis point decline in annual growth reported by the U.S. Bureau of Labor Statistics from 1996–2005 to 2006–2016.<sup>23</sup>

## 5 Conclusion

A large literature maintains that reallocation is central to growth in the U.S. and elsewhere. Although this appears particularly important in some cases, such as China's move from state-owned to private enterprises, we argue that it is not the main driver of growth in the U.S. in recent decades. We think that incumbent improvements of their own products, which necessitate relatively little reallocation of inputs, are the most important contributor to growth. As a result, the decline in entry and job reallocation rates was probably a minor contributor to the sharp slowdown in U.S. productivity growth in the last decade.

There are several caveats to our argument that we would like to stress. First, the threat of creative destruction may be a key factor motivating incumbents to improve their own products. See Aghion et al. (2005) for theory and evidence on this “escape from competition” motive. Thus creative destruction could have a larger, indirect effect on growth. Inversely, removing the threat of creative destruction might raise the return to own innovation so that they are strategic substitutes rather than complements.

---

<sup>23</sup>The BLS estimates that the rate of multifactor productivity growth (inclusive of the contribution of R&D and intellectual property and expressed in labor-augmenting terms) fell from 2.68% per year from 1996–2005 to 0.91% per year from 2006–2016.

Second, it is possible that knowledge spillovers are larger for entrants (and from creative destruction) than from own innovations. Akcigit and Kerr (forthcoming) present evidence that the patents of young firms are cited more frequently by other firms than are the patents of older firms (which receive more self-citations). They present a model in which own innovations run into diminishing returns and peter out on their own. In their setup, creative destruction is essential because it involves big jumps in quality that make further improvements by incumbents possible.

Third, our empirical implementation of a model-based growth decomposition is silent on optimal innovation policy. Atkeson and Burstein (2016) contend that the decentralized equilibrium is likely to produce too much creative destruction and too little own innovation because the former involves business stealing and the latter does not.

Finally, we do not answer the question of why productivity growth has fallen so steeply in the U.S. and elsewhere in the last decade. We do suggest, however, that the hunt should focus on whether incumbent firms are continuing to innovate successfully, as for example as in Bloom et al. (2017).

## References

- Acemoglu, Daron and James Robinson, *Why Nations Fail: Origins of Power, Poverty and Prosperity*, Crown Publishers (Random House), 2012.
- Acs, Zoltan J and Pamela Mueller, "Employment effects of business dynamics: Mice, gazelles and elephants," *Small Business Economics*, 2008, 30 (1), 85–100.
- Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and innovation: An inverted-U relationship," *The Quarterly Journal of Economics*, 2005, 120 (2), 701–728.

- Ahmad, Nadim and Eric Gonnard, "High growth enterprises and gazelles," in "International Consortium on Entrepreneurship (ICE) meeting. Copenhagen: ICE" 2007.
- Akcigit, Ufuk and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, forthcoming.
- Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 2014, 122 (5), 1013–1063.
- Atkeson, Andrew and Ariel T. Burstein, "Aggregate implications of innovation policy," 2016.
- Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, John Van Reenen et al., "The Fall of the Labor Share and the Rise of Superstar Firms," 2017.
- Bai, Chong-En, Chang-Tai Hsieh, and Zheng Michael Song, "The Long Shadow of a Fiscal Expansion," 2016.
- Baily, Martin Neil, Charles Hulten, and David Campbell, "Productivity Dynamics in Manufacturing Plants," *Brookings Papers: Microeconomics*, 1992, 4, 187–267.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-country differences in productivity: The role of allocation and selection," *The American Economic Review*, 2013, 103 (1), 305–334.
- Barth, Erling, Alex Bryson, James C Davis, and Richard Freeman, "Its where you work: Increases in the dispersion of earnings across establishments and individuals in the United States," *Journal of Labor Economics*, 2016, 34 (S2), S67–S97.
- Bils, Mark, Peter J. Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," 2017.
- Birch, David L, "Who creates jobs?," *The public interest*, 1981, (65), 3.
- Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," 2017.
- Bollard, Albert, Peter J. Klenow, and Huiyu Li, "Entry costs rise with development," 2016.

Davis, Steven J, John C Haltiwanger, and Scott Schuh, “Job Creation and Destruction,” *MIT Press Books*, 1998.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, “The Secular Decline in Business Dynamism in the US,” *Manuscript, University of Maryland*, 2014.

—, —, —, and —, “Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown,” *American Economic Review*, 2017, (107(5)), 322–326.

—, —, Ron S. Jarmin, and Javier Miranda, “Changing Business Dynamism and Productivity: Shocks vs. Responsiveness,” 2017.

Foster, Lucia, John C Haltiwanger, and C.J. Krizan, “Aggregate Productivity Growth. Lessons from Microeconomic Evidence,” *New Developments in Productivity Analysis*, 2001, pp. 303–372.

—, John Haltiwanger, and Chad Syverson, “Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?,” *The American economic review*, 2008, 98 (1), 394–425.

—, —, and —, “The slow growth of new plants: Learning about demand?,” *Economica*, 2016, 83 (329), 91–129.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow, “How Destructive is Innovation?,” 2016.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez, “Capital allocation and productivity in South Europe,” *The Quarterly Journal of Economics*, 2017, p. forthcoming.

Gourio, Francois and Leena Rudanko, “Customer capital,” *Review of Economic Studies*, 2014, 81 (3).

Haltiwanger, John, Robert Kulick, and Chad Syverson, “Misallocation Measures: Glowing Like the Metal on the Edge of a Knife,” 2016.

Hopenhayn, Hugo A, “Firms, misallocation, and aggregate productivity: A review,” *Annual Reviews in Economics*, 2014, 6 (1), 735–770.

- Hottman, Colin J, Stephen J Redding, and David E Weinstein, “Quantifying the sources of firm heterogeneity,” *The Quarterly Journal of Economics*, 2016, 131 (3), 1291–1364.
- Hsieh, Chang-Tai and Enrico Moretti, “Housing Constraints and Spatial Misallocation,” *NBER Working Paper Series*, 2017, w21154.
- and Peter J Klenow, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 2009, 124 (4), 1403–1448.
- and Peter J. Klenow, “The Life Cycle of Plants in India and Mexico,” *Quarterly Journal of Economics*, 2014, 129 (3), 1035–1084.
- Kehrig, Matthias, “The Cyclical Nature of the Productivity Distribution,” 2015.
- Klette, Tor Jakob and Samuel Kortum, “Innovating firms and aggregate innovation,” *Journal of Political Economy*, 2004, 112 (5), 986–1018.
- Olley, G. Steven and Ariel Pakes, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 1996, 64 (6), 1263–1297.
- Peters, Michael, “Heterogeneous Mark-ups, Growth and Endogenous Misallocation,” 2016.
- Restuccia, Diego and Richard Rogerson, “The Causes and Costs of Misallocation,” *Journal of Economic Perspectives*, August 2017, 31 (3), 151–74.
- Sato, Kazuo, “The Ideal Log-Change Index Number,” *The Review of Economics and Statistics*, 1976, pp. 223–228.
- Schumpeter, Joseph A., *Business Cycles*, Cambridge University Press, 1939.
- Syverson, Chad, “What determines productivity?,” *Journal of Economic literature*, 2011, 49 (2), 326–365.
- Vartia, Yrjö O, “Ideal Log-Change Index Numbers,” *Scandinavian Journal of Statistics*, 1976, pp. 121–126.