Comments on “The Reallocation Myth” by Chang-Tai Hsieh and Peter Klenow

Remarks by John Haltiwanger, University of Maryland[1]

Hsieh and Klenow (2017) (hereafter HK (2017)) challenge what they characterize as conventional wisdom about the role of reallocation in productivity growth and the contribution of the decline in the pace of business dynamism for the post 2000 decline in productivity growth in the U.S. In a nutshell, the arguments of the paper are as follows. First, they argue that their findings suggest there has not been an increase in allocative efficiency in the U.S. over the last several decades. Their evidence on this uses the framework they developed in their highly influential paper (Hsieh and Klenow (2009 -- hereafter HK (2009)) to measure allocative efficiency. They draw their inferences using U.S. manufacturing data over time. They argue that the lack of evidence of an increase in allocative efficiency raises questions about the role of reallocation in productivity growth. Second, they consider the potential role of reallocation as capturing the contribution of creative destruction in models of endogenous innovation. They calibrate a simple model of endogenous innovation to match moments of job creation and destruction for the U.S. private sector along with moments of aggregate productivity growth. Through the lens of that model, they argue that about 25% of productivity growth is due to creative destruction with most of the productivity growth occurring through within firm improvements in product quality. In turn, they note that there has been a decline in job reallocation and particularly a decline in the contribution of entrants. Their calculations suggest that this decline in dynamism makes a non-trivial contribution to the decline of productivity growth (as much as 10%) but most remains unexplained.

In these comments, I review the arguments underlying these inferences in detail paying attention to issues of model specification, empirical measurement and identification and their comments about alternative methods that have been used that suggest different conclusions. These details are important and disagreements about these issues are part of the normal course of scientific debate. Before proceeding to these details, I first provide some remarks about broad issues that are missing or neglected in HK (2017). First, HK (2017) base their inferences on models that omit key elements – referred to as “reallocation frictions” – that characterizes much of the research to which they are responding. I review both direct and indirect evidence related

[1] These comments have benefitted from helpful conversations with Steven Davis and Chad Syverson.

to such frictions that HK (2017) do not confront. Second, the quantitative model results on which HK (2017) draw depend critically on calibration choices related to young firm activity. I briefly review key empirical results from the existing literature suggesting that the model described by HK (2017) biases results against a more important role for young firms. Part of this reflects the neglect of reallocation frictions. In the HK (2017) model and calibration, young firms that make major innovations should immediately gain large market shares. But reallocation frictions imply these dynamics likely play out over a much longer period than they allow. Finally, the creative destruction models that HK (2017) calibrate are more likely relevant for innovative intensive industries. Ignoring sectoral variation in the nature and type of innovation is likely important for understanding the contribution of different types of reallocation to productivity. These broad issues are relevant for the discussion of details that follow and taken together with the technical details provide ample reason to be skeptical about the conclusions in HK (2017).

I. Broad Issues

A. The Importance of Reallocation Frictions

The first broad concern is a general neglect of the role of reallocation frictions that influence and impede the pace of reallocation. HK (2017) mention such frictions but critically they neglect the potential role of changes in these reallocation frictions over time. It is potential changes in the latter that underlie the hypothesis relating changing business dynamism and productivity growth.

Before proceeding to discussing changes in these frictions, first consider what is meant by such frictions. These reallocation frictions include costs to business entry, to post entry growth, and to business exit and contraction. Some aspects of these frictions are an inherent part of the costs of doing business. It takes time and resources for entrepreneurs to start up businesses. Recruiting, search, hiring and training costs are an inherent part of firms adding new workers. Like starting a business, it takes time and resources to acquire and install new capital whether it is physical capital or organizational capital in terms of how a business is organized. Investment in any form of capital (physical, organizational or human) often has firm-specific components that imply "irreversibilities". Developing customer bases is another friction that young businesses face, as do incumbents entering new markets in terms of products or spatial expansion. The latter includes entering new markets globally. While some aspects of these reallocation frictions are inevitable, these costs can also vary over time, industry and country based on the business climate in terms of market institutions and the regulatory environment.

How are these issues important in this context? It is well established theoretically and empirically that an increase in reallocation frictions will reduce the pace of job reallocation (and capital and output reallocation) and lower productivity. Hopenhayn and Rogerson (1993)'s seminal paper on the impact of an increase in reallocation frictions on the pace of job turnover and productivity is one example. HK (2017) don't cite or discuss Hopenhayn and Rogerson (1993) or most of the large accompanying subsequent empirical and theoretical literature. [2] Instead, HK (2017) state in footnote 8 that "adjustment costs actually affect the level of aggregate productivity, not its growth rate." The latter is technically correct in the following sense. An economy that experiences a once and for all increase in adjustment costs impeding reallocation will in the long run experience a decline in in level of reallocation and productivity. HK (2017) relegate the role of reallocation frictions to this perspective that they only influence the level of productivity.

However, this is a narrow perspective for two reasons. First, the transition path for even a once and for all change a reallocation frictions can be long. Davis and Haltiwanger (1999) illustrate this point by calibrating a model with one time changes in reallocation frictions. Second, changes in reallocation frictions in an industry or country may exhibit secular trends. Taken together, these points provide the underpinnings for the hypothesis that an increase in reallocation frictions may underlie the post 2000 period decline in business dynamism and in turn have been a drag on medium-term productivity growth over this period. Is there any evidence in favor of this hypothesis? The answer is yes both indirectly and directly. Indirectly, Decker et. al. (2017) show that the responsiveness of firms to idiosyncratic productivity shocks has declined over this period, particularly but not exclusively among young firms. That is, firms with high productivity realizations are less likely to grow and firms with low productivity realizations are less likely to contract and exit. This is the pattern one would expect following an increase in reallocation frictions in the Hopenhayn and Rogerson (1993) framework. Moreover, the Decker et. al. (2017) evidence suggests this is not a once and for all abrupt change but rather

[2] There is a large accompanying literature on measuring and analyzing adjustment costs at the micro level including Cooper and Haltiwanger (2006), Cooper, Haltiwanger and Willis (2007) and Elsby and Michaels (2013).

a secular trend in the declining responsiveness of businesses to shocks. This may be the result of transition dynamics or a secular increase in reallocation frictions.

Direct evidence of what is driving an increase in reallocation frictions is still under investigation but Davis and Haltiwanger (2014) discuss several possibilities and offer some evidence. For example, they highlight the findings of Autor et. al. (2006, 2007) that show that through legal precedents there has been a decline in the employment at will doctrine in the U.S. Davis and Haltiwanger (2014) present evidence that this change in the institutional environment yields a non-trivial decline in the pace of job reallocation. They argue that the Autor et. al. evidence is symptomatic of increasing frictions that employers face in hiring and firing workers. For example, Krueger and Kleiner (2013) show that there has been a dramatic increase in the fraction of jobs requiring an occupational license (from roughly 5% a few decades ago to over 30% currently) where many of these requirements are state-specific (so this impedes both occupational and geographic mobility). The consequence of these changes is not yet fully understood but they also discuss possible increased frictions in financial markets and product markets that may increasingly be impeding business entry and reallocation.

HK (2017) mostly neglect this line of argument and inquiry and thus the substantial evidence suggesting this is likely an important contributor to the sluggish productivity growth post 2000. Moreover, this neglect yields a different perspective on some of the claims in HK (2017). The increasing reallocation frictions hypothesis implies that we should be observing a rise in dispersion in revenue productivity measures and a decline in allocative efficiency. HK (2017) acknowledge the former is present in the data especially post 2000 but not the latter where they argue there has not been much change. As discussed in the details below, there are inconsistencies in the arguments and evidence that HK (2017) use on these points.

B. Measuring the Contribution of Young Firms to Innovation

A second broad area where the HK (2017) approach and evidence potentially has important limitations is their calibration of the role of young firms in being disproportionately important in major innovations. The creative destruction model that HK (2017) calibrate does permit entrants to differ from incumbents in the probabilities of innovation and the arrival rates of new varieties. However, the model parameters are primarily determined by the shape of the job creation and destruction rates across employment growth rate bins by firm age. There is considerable evidence they neglect in their calibration that suggest entrants and young businesses exhibit additional important differences from incumbents. Some of this HK (2017) acknowledge in their concluding remarks (e.g., the work of Akcigit and Kerr (2017)) but this is only part of the relevant evidence. Acemoglu et. al. (2013) develop a related model of creative destruction where young firms have a much higher likelihood of making major innovations. They motivate and discipline this model with evidence that young firms are, in the innovative intensive industries, more R&D and patent intensive. Moreover, young firms' exhibit much more skewed growth rate distributions of employment and sales than mature firms. It is far from clear that the HK (2017) calibrated parameters can match these patterns in the data. In the discussion of details below, there are additional issues that raise questions about the interpretation of the HK (2017) calibration of creative destruction.

C. Reallocation Frictions and Innovation from Young Firms Interact

These two broad areas of limitation may interact in important ways. The underlying Garcia-Macia, Hsieh and Klenow (2016) (hereafter GHK (2016)) paper that HK (2017) base their creative destruction analysis seems to acknowledge these limitations more than HK (2017). In the concluding remarks of GHK (2016), they state (on page 33):

We assumed no frictions in employment growth or misallocation of labor across firms. In reality, the market share of young firms could be suppressed by adjustment costs, financing frictions, and uncertainty. On top of adjustment costs for capital and labor, firms may take awhile to build up a customer base, as in work by Gourio and Rudanko (2014) and Foster et al. (2016). Irreversibilities could combine with uncertainty about the firm's quality to keep young firms small, as in the Jovanovic (1982) model. We defined young firms as those younger than five years, but these dynamics could play out for longer.

All of these remarks fit into reallocation frictions discussed above. All will limit the employment share of entrants (a key moment GHK (2016) and HK (2017)) so that inferring the contribution of innovation by entrants by this moment may be quite misleading. Note that many of these factors imply what the last sentence above suggests – that it will take longer for the dynamics of young firms' contribution to play out. This will imply that more of the contribution of young firms to creative destruction is captured by incumbents (more than five years old). Moreover, as discussed above, increased reallocation frictions may further limit the share of entrants and in turn yield greater limitations of inferring the contribution of young firms to innovation.

In short, the neglect of reallocation frictions in their calibration of their creative destruction model of innovation biases the result in favor of the finding that own incumbents innovation is the primary contributing factor to productivity growth. This point provides perspective on a seemingly appealing aspect of the HK (2017) arguments and identification. They argue that if creative destruction is the dominant force then the small changes in job creation by mature incumbents should not be so important. However, reallocation frictions will slow down and dampen the dynamics of employment reallocation induced by creative destruction implying that firms (young or mature) making major innovations will only slowly gain market share.

Specific examples help make the case that five years is not enough for innovation by new firms to develop their products, processes and business model. Amazon started in 1994 and by 1999 had grown to 7600 employees.[3] Under the HK (2017) calibration this would be a sizeable contribution to job creation from entrants and contribute substantially to creative destruction from innovation. However, it was clear in 1999 that Amazon was still very much in its formative stages in terms of developing its business model, its organizational structure and its customer base. Indeed, there was much skepticism that the Amazon business model would even survive in the early 2000s. In 2017, Amazon has over 300,000 employees and is the largest internet company in the world by revenue. All of the growth since 1999 would be attributed to mature incumbents in the HK (2017). It is only since 2011 that it would surpass the 20% average annual employment growth rate over a five year horizon that would make it a "gazelle" in HK (2017). And in many ways, it is only now that Amazon is beginning to fundamentally transform retail trade. It also does not seem that Amazon's post 5-year old growth is well described as an incumbent slowly making own innovations adding product varieties as in HK (2017).

Amazon is only one example but over the very long run innovation is also tightly linked to reallocation. For example, only 60 firms in today's S&P 500 were there at the index's inception 60 years ago; similarly only GE is left from the original Dow Industrials. The vast majority of economic activity today is being conducted by firms that (even accounting for mergers etc.) did not exist 100 years ago.

D. The Nature of Reallocation and Innovation Varies Substantially Across Sectors

The creative destruction model of innovation that HK (2017) apply to the entire U.S. private sector is better suited to some sectors than others. Acemoglu et. al. (2013) intentionally

[3] All of the information reported here is from publicly available information such as 10-K filings.

focus their analysis to innovative intensive sectors where patents and R&D are concentrated along with major innovations that fit well into this type of model. This includes the ICT sectors of the economy that Fernald (2014) and Byrne et. al. (2016) show are the primary drivers of the surge in productivity in the 1990s and the post 2000 decline.

Applying this model to all sectors neglects the heterogeneity in the nature of entrepreneurs emphasized by Hurst and Pugsley (2017). Hurst and Puglsey (2017) emphasize that some sectors (e.g., dry cleaners in the service sector) are dominated by business owners whose ownership motive is non-pecuniary benefits of, for example, being one's own boss. In such sectors, the shape of the job creation and destruction distribution is likely dominated by small changes as such businesses are buffeted by idiosyncratic shocks to their location and local economy. Alternatively, in sectors like retail trade, the reallocation that occurred in the 1990s and 2000s primarily reflected a change in the business model from small, single unit establishment firms to large, national chains with many locations (see, e.g., Foster, Haltiwanger and Krizan (2006)). Firms like Wal-Mart were able to take advantage of IT and globalization to build efficient distribution networks and global supply chains. This has been productivity enhancing reallocation that would not be well captured by using the HK (2017) model calibrated to the shape of the job creation and destruction distribution. And as noted above, retail trade appears to be at the cusp of another form of transformative reallocation with Amazon – a 1990s era startup that has taken a long time to reach this transformative stage.

II. Technical Details

I now turn to discussing some of the technical details that raise further questions about the interpretation of the evidence in HK (2017). This discussion is more relevant for specialists in the academic literature since much of it involves model and empirical specification and measurement details.

A. *Statistical Accounting vs. Structural Decompositions of Aggregate Productivity Growth*

HK (2017) provide an overview of some of the accounting decompositions and their results. It is not clear why they focus primarily on the accounting decomposition developed by Baily, Hulten and Campbell (1992) (BHC). While the latter is a path breaking paper in the literature on firm productivity dynamics, it is now well understood since at least Griliches and Regev (1995) (GR) and Foster, Haltiwanger and Krizan (2001) (FHK) that the specific decomposition BHC used has significant limitations even within the context of statistical accounting decompositions.[4] The particular specification of the BHC decomposition has not been actively used in the literature since the GR and FHK papers.

However, the main concern of HK (2017) is that it is difficult to interpret the BHC/FHK/GR statistical accounting decompositions without an accompanying model of heterogeneous firm dynamics. This is a fair point and has been readily acknowledged in the literature in a number of ways. First, an alternative statistical accounting decomposition developed by Olley and Pakes (1996) has proven easier to map to theory. This is a statistical approach that decomposes a weighted mean into an unweighted mean and a covariance term reflecting the covariance between a firm's market share and productivity.[5] Is this a moment useful for identifying changes in distortions and frictions? A number of heterogeneous agent models suggest that it is an informative moment. Bartelsman et. al (2013) show that the OP covariance term (measured using technical efficiency or what has become known as TFPQ) declines with a rise in the dispersion of idiosyncratic distortions in a heterogeneous firms equilibrium model with overhead labor and quasi-fixed capital. This is intuitive since a greater dispersion in distortions implies that the covariance between TFPQ and firm size will be diminished. In a related manner, Decker et. al. (2017) show that in a calibrated heterogeneous firms model with adjustment costs that an increase in adjustment frictions leads to a decline in OP covariance terms defined using TFPQ. This latter finding is consistent with Hopenhayn and Rogerson's finding that an increase in reallocation frictions reduces productivity via a decline in allocative efficiency.

There is an important conceptual and measurement issue about what empirical measures of OP covariances are relevant for testing hypotheses or calibrating and estimating models. The issue is that the theory implies changes in OP covariances using measures of TFPQ at the firm level. However, measuring TFPQ empirically is a challenge and much of the literature has used revenue productivity measures as proxies for TFPQ. Specifically, researchers have often used TFPR. The latter is formally defined in Foster, Haltiwanger and Syverson (2008) as P*TFPQ

---

[4] BHC/FHK/GR are within/between decompositions of changes in a weighted mean. Such within/between or shift/share decompositions have been used in many settings to provide perspective on the contribution of changing composition—i.e., reallocation. BHC extended the decomposition to incorporate entry/exit but without taking into account that the interpretation of the balanced panel between term does not carry over to an unbalanced panel without modification. See FHK for details.

[5] Important modifications of the Olley and Pakes (1996) decomposition can be found in FHK and more recently in Melitz and Polanec (2015). The latter paper nicely extends the decomposition into a dynamic setting breaking out the contribution of entry and exit.

where P is the plant-level price for the product of a given plant. The relationship between TFPQ and TFPR plays a critical role in this literature including in the inferences made in HK (2017).

I discuss in detail the issues of what we know about the relationship between TFPQ and TFPR (and other revenue productivity measures) in the next section but it is useful to elaborate more now on why it is so important in this context. HK (2017) discuss the Olley and Pakes (1996) decomposition as though the approach intentionally focuses on using a weighted mean of revenue productivity (they mention revenue labor productivity) in the decomposition. This is misleading since this statistical decomposition (as well as the others) is intended in principle to be applied using TFPQ. In the HK (2009) framework, an increase in the distribution of idiosyncratic distortions will lower the OP covariance defined using TFPQ (with output weights) as discussed above (and demonstrated in Bartelsman et. al. (2013)).[6] The latter paper shows that this same inference carries over to OP covariances computed using TFPR or revenue labor productivity in the presence of overhead labor and quasi-fixed capital. Decker et. al. (2017) show that an increase in adjustment frictions will yield a lower OP covariance using either TFPQ, TFPR or revenue labor productivity in the presence of a baseline of adjustment frictions.

A key issue is what is the benchmark and the impact of this benchmark on the relationship between TFPQ and revenue productivity measures. HK (2017) discussion of the OP covariance using revenue labor productivity uses the frictionless/distortionless efficient benchmark where dispersion in revenue labor productivity is zero so that the correlation between TFPQ and these measures is zero. But starting from a benchmark with either overhead labor or adjustment costs (or other departures from the HK (2009) strong identifying assumptions discussed below), TFPR and revenue labor productivity exhibit dispersion in the absence of distortions and are positively correlated with TFPQ. From such alternative (and more empirically relevant) benchmarks, the OP covariance terms using either TFPR or revenue labor productivity decline with increases in either the distribution of distortions or adjustment costs.[7]

The frictionless/distortionless benchmark is problematic for a number of reasons. Many of the reallocation frictions highlighted above reflect frictions the social planner cannot avoid. Moreover, the frictionless benchmark has implausible empirical properties for the patterns of firm-level employment growth. Decker et. al. (2017) show that if one calibrates a model of firm

---

[6] This implication holds even in the HK (2009) framework without any adjustment costs or overhead labor.
[7] These statements hold based on benchmarks of overhead costs and adjustment costs calibrated to U.S. data.

dynamics with zero adjustment frictions with standard parameters for the curvature of the revenue function and the shock processes, that the implied pace of annual job reallocation is several times larger than the actual pace of annual job reallocation. It is apparent that reallocation frictions are needed to account for the observed pace of reallocation.

Before turning to the discussion of details on TFPQ and TPFR, it is useful to review what we have learned from an alternative structural decomposition of aggregate productivity growth that is based on economic theory. Petrin and Levinsohn (2012) develop a structural decomposition of aggregate productivity growth (APG) from economic theory. Their approach explicitly takes into account optimizing firms and the frictions that such firms face. The structural decomposition has multiple terms – a within firm productivity growth term, reallocation terms and the role of changing fixed costs of operation. The reallocation terms derived from economic theory have the intuitive interpretation that firm-level input growth contributes to APG only if it is moving resources from firms that have lower value of marginal products to firms with higher value of marginal products. As such, these decompositions overcome a number of the concerns that HK (2017) raise about the statistical decompositions. For example, these decompositions are consistent with the statements in HK (2017) about reallocation and marginal revenue products.

As in HK (2009) and HK (2017), there are challenges in implementing these structural decompositions because they rely on estimates of the key parameters of the structure of demand (such as demand elasticities) and output elasticities.[8] Petrin, White and Reiter (2011) (PWR) and more recently Foster, Grim, Haltiwanger and Wolf (2017) (FGHW) have implemented these structural decompositions of APG for U.S. manufacturing plants. The FGHW paper's focus is more on the challenges of identifying the key parameters and the sensitivity of the contribution of different terms in the decomposition to these estimates. These details are important but for the current purpose two messages from these papers are of critical importance. Both PWR and FGHW find that on average the reallocation contribution to APG tends to exceed the within plant contribution to APG. PWR estimate that more than 50% of APG in US manufacturing productivity growth is due to reallocation. FGHW show that estimate might be as low as 25%

[8] Implementation requires these estimates but the decomposition itself imposes much less restrictions on the data than HK (2009) and HK (2017). That is, the decomposition is consistent with permitting non CRTS, adjustment costs, fixed costs and non CES demand.

across estimation methods but this is still a substantial contribution. In addition, the relatively low contribution of within plant productivity growth that PWR and FGHW find is relevant for HK (2017) model of creative destruction. I return to this point below but the main point is that HK (2017) neglect entirely the insights and findings from the structural decomposition literature. The theoretical underpinnings of these decompositions overcome key objections they have about the statistical accounting decompositions. Moreover, all results show an important role for reallocation in productivity growth.

Before turning to further discussion of TFPQ vs. TFPR details, a few final remarks about statistical accounting decompositions vs. structural decompositions are in order. The advantage of the structural decomposition developed by PL is that it isolates reallocation that is moving resources from low value of marginal product activities to high value of marginal product activities. All of the statistical decompositions (BHC/GR/FHK and OP) using TFPQ (or proxies) neglect this important marginal distinction. This is not a limitation under CRTS production technologies and perfect competition as in this case it is optimal for resources to be allocated to the highest TFPQ firms within an industry in the absence of distortions or frictions.[9] This is because in this case TFPQ, TFPR and the marginal revenue product of the (composite) input are perfectly correlated under these assumptions. The presence of adjustment costs prevents the allocation of activity to the highest TFPQ firms and yields dispersion in TFPQ in equilibrium under the assumptions of CRTS and perfect competition. It is, for example, this perspective that underlies the productivity dispersion in canonical search and matching models of the labor market such as Mortensen and Pissarides (1994). In that model, dispersion in productivity arises from frictions in the labor market and in the absence of frictions all resources are optimally allocated to the most productive technology. Caballero and Hammour (1994, 1996) models of creative destruction share this feature.[10]

---

[9] All of the statistical accounting decompositions are based on defining an industry-level productivity index as the weighted average of firm level productivity (for this discussion we discuss this in terms of TFPQ). The standard industry-level measure of productivity is total output per unit (composite) input. The weighted average of firm-level productivity is equivalent to the standard measure under CRTS and perfect competition (using appropriate weights where the latter is the ratio of the firm-level composite input using output elasticities to industry-level composite input using the same output elasticities). Departures from these assumptions imply the weighted average index will deviate from the standard measure for reasons related to this discussion. In practice, the weighted average measures track standard measures reasonably well. But appropriate caution is needed in the use of these weighted average indices (and in turn the statistical decompositions) with departures from CRTS and perfect competition.

[10] The insights of these models don't depend on perfect competition and CRTS. However, the implication in these models that all resources should be allocated to the highest TFPQ production units in the absence of adjustment

TFPQ vs. TFPR?

The measurement and interpretation of TFPQ and TFPR is critically important for the inferences in HK (2017) both in terms of their comments about the literature but also for the findings they present on allocative efficiency. The reason is that HK (2017) base their measurement of distortions, TFPQ, and allocative efficiency on the HK (2009) model that makes very strong identifying assumptions for these key measures. HK (2009) demonstrate that under the assumptions of iso-elastic demand, constant returns to scale in a Cobb-Douglas production technology, no fixed costs and no adjustment costs that dispersion in TFPR should be zero.[11] The logic is that under these assumptions establishments will equate marginal revenue products to common factor prices. Under these assumptions, dispersion in TFPQ will yield an equilibrium size distribution but no dispersion in TFPR. Establishments with higher TFPQ will use more inputs, have higher output and charge lower prices. A further implication of this framework, as emphasized by Haltiwanger, Kulick and Syverson (2017) (HKS), is that under these assumptions establishment-level prices have an inverse unit elastic relationship with respect to TFPQ. That is, plants with higher TFPQ will be larger with lower prices where the latter will decline in exactly the right proportion so that TFPR is invariant.

HK (2009) use this insight to argue that observed dispersion in TFPR must reflect some distortion preventing equalization of marginal revenue products. This is the critical identifying assumption that HK (2017) use to identify what they denote as idiosyncratic distortions on page 9 of their paper. In turn, HK (2009) and HK (2017) use this structure to identify TFPQ which they cannot measure directly since in the datasets they use they don't observe establishment level prices. Their model implies that plant-level TFPQ can be measured (up to a factor of proportionality at the industry level) as a transformation of plant-level revenue.[12] This yields an implied measure of plant-level output that permits them to measure TFPQ as output per unit input.[13]

costs does depend on these assumptions. For models with adjustment frictions and decreasing returns see Cooper, Haltiwanger and Willis (2007) and Elsby and Michaels (2013).

[11] They also require that the demand and output elasticities be common to all plants in the same industry.

[12] In practice, this is done with a calibrated and not estimated CES parameter they assume is the same across all industries, countries, and time.

[13] In practice, HK (2009) measure the composite input for each plant assuming the same output elasticities for factors such as capital and labor across all industries, countries and time.

Given this background, what do we know about the relationship between TFPR and TFPQ both theoretically and empirically? Theoretically, the identifying assumptions made by HK (2009) have a knife-edge quality. Haltiwanger, Kulick and Syverson (2017) (HKS) show that the only demand and technology structure that permit the HK (2009) identification is CES demand with Cobb-Douglas CRTS technology. Any departures from these assumptions and TFPR exhibits dispersion that is correlated with fundamentals such as TFPQ and demand shocks even in the absence of distortions. Asker et. al. (2014) and Decker et. al. (2017) show that TFPR will exhibit dispersion and be highly correlated with TFPQ in any standard model of adjustment frictions.[14] Bartelsman et. al. (2013) show that TFPQ and TFPR will be highly correlated in a model with overhead labor and quasi-fixed capital.

Turning to empirics, for selected products in the US, Foster, Haltiwanger and Syverson (FHS) (2008, 2016) use price and quantity data that enables direct measurement of TFPQ in a manner that also enables consistent measurement of TFPR=P*TFPQ. They find that the correlation between TFPR and TFPQ is about 0.75. Eslava et. al. (2013) use price and quantity data for Colombia using a similar methodology to FHS and find similarly high correlations. Decker et. al. (2017) and FGHW (2017) show that an alternative revenue productivity measure from revenue function estimation is a function of fundamentals even under the HK (2009) assumptions. They find that this measure of fundamentals is highly correlated with TFPR (again about 0.75 in Decker et. al. (2017)).[15] The latter holds for the entire US manufacturing sector (not just the selected products in FHS).

HK (2017) report that the correlation between TFPR and TFPQ using their indirect method for estimating TFPQ is low in U.S. manufacturing. They use this finding to suggest that any paper that uses TFPR as a proxy for TFPQ is problematic. This is based on the argument that in their model TFPR reflects distortions and not TFPQ and that empirically the relationship between TFPQ and TFPR is weak. But this weak relationship is an inference that is dependent

[14] Gopinath et. al. (2017) include capital adjustment costs and note that it generates dispersion in TFPR in the absence of distortions. They report that the correlation between their revenue function residual (which is a function of demand and TFPQ) is highly correlated with the indirect measure of TFPQ they generate using the HK (2009) methodology.

[15] FGHW find slightly lower correlations but they push the proxy method harder than Decker et. al. (2017) by estimating the revenue elasticities at the 6-digit NAICS level while Decker et. al. (2017) estimate at the 3-digit level. FGHW show that the proxy method estimates are noisier at a higher level of disaggregation. These are methods that are sensitive to outliers that are more important in smaller industry cells.

on the model, identifying assumptions and on the output elasticities and demand elasticities that HK (2009) and (2017) use in their calibrations.

The difference in TFPQ and TFPR correlations that HK (2009) find relative to the rest of the literature raises questions about the indirect method that HK (2009) and (2017) use to measure TFPQ. HKS provide additional evidence that raises further questions. HKS use the same data as FHS with direct measures of TFPR and TFPQ and they also compute the indirect method based TFPQ (what they call TFPQ_HK) using the same transformation HK (2009, 2017) use in their analysis. In doing so, they use detailed product specific estimates of output elasticities and demand elasticities (that vary considerably across products). They find a weak correlation between directly measured TFPQ and TFPQ_HK (less than 0.10). They also find that TFPQ_HK exhibits much higher dispersion than TFPQ. They also test the inverse unit elastic implication of the HK (2009) identifying assumptions. They soundly reject this hypothesis finding an elasticity of prices with respect to TFPQ of about -0.5. HKS also conduct external validity tests of these alternative measures. They find (like others in the literature have) that plants with high measured TFPR and TFPQ are much more likely to survive. If a high TFPR is supposed to be a plant facing a high distortion, it is not clear why the most distorted plants should be the most likely to survive. They also find that high TFPQ_HK plants only have a weak relationship with survival or in some specifications are less likely to survive. This in inconsistent with the evidence for the directly measured TFPQ as well as economic theory.

Of course the HKS results apply to a particular sample, one that is by no means representative of all production settings. However, there is evidence from other parts of the literature that suggest these findings are more general. For example, there is a large and separate empirical literature on pass-through rates which suggest that the result that prices don't fully respond to TFPQ may be typical. These studies all find in diverse market settings that pass through of costs into prices is less than one-for-one (see HKS for more discussion and citations). More broadly, numerous studies have used TFPR measures and related them to favorable outcomes such as growth and survival (see, e.g., Syverson (2011) for many examples) as well as to determinants of differences in firm performance such as management practices. For example, Bloom et. al. (2017) find that establishments with more structured management practices have much higher measured TFPR. These findings re-inforce the interpretation that TFPR is related to favorable outcomes and indicators of business performance, rather than being an indicator of distortions.

In sum, the identifying assumptions for measuring distortions through TFPR and for identifying TFPQ rest on knife edge assumptions on the structure of demand, technology, the absence of fixed costs and the absence of adjustment costs. In addition, the empirical evidence suggests that these assumptions are not supported by the evidence. This raises substantial questions about whether HK (2009, 2017) have identified distortions via TFPR or TFPQ through their indirect method of transforming revenue using a demand elasticity. Their measures of allocative efficiency depend critically on their identification of distortions and TFPQ. As such, it is difficult to interpret their measures of allocative efficiency as well as their comments about the use of TFPR and related revenue productivity measures by others in the literature.

Some of the qualitative inferences about changes in the distribution of distortions in HK (2009, 2017) are likely robust to these specification error issues. Even if dispersion in TFPR likely reflects many other factors than distortions, it is still likely that an increase in distortions should lead to an increase in dispersion in TFPR.[16] The identification problem is to isolate variation in the data where the difference in dispersion in TFPR is due to difference in distortions rather than differences in these many other factors that may be at work. This is akin to standard omitted variables problems confronting identification. In this case, these omitted variables problems appear to be severe. This identification problem is arguably much greater in cross country comparisons than within country variation over time. A good example of using within country variation over time is Gopinath et. al. (2017).[17]

Even if one isolates variation in dispersion in TFPR that might be due to changes in distortions, there is alternative and related hypothesis is that a rise in dispersion in TFPR is due to an increase in reallocation frictions. Of course, one might want to lump idiosyncratic distortions and reallocation frictions together but the discussion in these comments emphasizes that reallocation frictions have dynamic implications that “black box” distortions do not. For example, a rise in reallocation frictions reduces the responsiveness of firms to shocks, reduces

[16] For example, Bartelsman et. al. (2013) show that an increase in idiosyncratic distortions tends to raise dispersion in TFPR even in a model with fixed (overhead labor) costs and quasi-fixed capital.
[17] Gopinath et. al. (2017) go much further than just exploiting within country variation over time. They develop a heterogeneous agents model with capital adjustment costs and financial frictions to explore the evolution of micro and macro TFP in Southern Europe during the financial crisis.

the pace of job reallocation and increases the correlation between TFPQ and TFPR. These additional predictions provide additional moments to help identify the role of changing reallocation frictions.

Returning to the evidence in HK (2017), they note that several papers have found evidence of rising TFPR (and closely related revenue labor productivity) dispersion in the US. Through the lens of either the increase in distortions or reallocation frictions interpretation, this implies a decline in allocative efficiency and thus productivity. Under the rising frictions interpretation, this also is consistent with declining business dynamism. Thus, while they don't interpret it in this manner, this is evidence presented in HK (2017) consistent with declining business dynamism being associated with the decline in productivity.

This latter discussion highlights some tension in HK (2017) over rising TFPR and implications for declining allocative efficiency. They note that Bils, Klenow and Ruane (2017) suggest the rising dispersion in TFPR in the U.S. may be due to rising measurement error in the Annual Survey of Manufactures. But as they also note, administrative data sources also show rising dispersion making the measurement error explanation less plausible.

B. *Calibration of the Creative Destruction Model*

Two key challenges for interpreting the HK (2017) calibration of the creative destruction model have already been noted. First, the calibration neglects evidence that young firms have a disproportionate likelihood of making major innovations. Second, the calibration neglects any reallocation frictions that may imply that entrants/young firms' contribution is restricted in early years so that their ultimate contribution is improperly assigned to older incumbent firms.

There are some other details of the calibration in HK (2017) that also raise questions of interpretation. HK (2017) interpret their evidence that it is slow growing, mature incumbents that account for most productivity growth. This inference is based on Table 2 where incumbents (older than five) with annual growth rates of less than 20% account for 37% of job creation and 65% of aggregate growth. However, this cutoff of 20% is the annual average over a five year horizon. Firms that grow 20% on average annually for a five year period are exhibiting phenomenal growth. This implies that firms with high growth rates (e.g., 90% cumulatively over a five year horizon) are in the "slow growth" category. For example, I suspect if 5% average annual growth rates over a five year period were used as the cutoff that the inferences in Table 2 would be quite different. A 5% growth cutoff is still a rate that exceeds the annual average growth rate of firms. More generally, many calibration choices made by HK (2017) and GHK (2017) bias the results against a more important role for entrants and creative destruction.

Finally, I return to the evidence discussed above about the LP/PWR decompositions. These decompositions find a limited role for within plant productivity growth in U.S. manufacturing. If it is the large, mature incumbents making own improvements in quality contributing the most, we should be observing more evidence of within plant productivity growth. The within plant productivity growth evidence is restricted to the manufacturing sector which is a limitation. More broadly, this evidence is a reminder that GHK (2016) and HK (2017) are making inferences about within firm productivity growth without using any direct information about within firm productivity growth at the micro level. Measuring within firm productivity growth at the micro level is a challenge especially moving beyond the manufacturing sector. However, until there are robust within firm productivity growth measures available by firm age, appropriate caution is required in any inferences drawn about the role of slow growing, mature incumbents.

III. Concluding Remarks

What do we learn from HK (2017)? Based upon the above discussion:

- HK (2017) raise concerns about linking statistical accounting decompositions to economic theory. This is a valid concern but one that has already been addressed in the literature. Increasingly the Olley Pakes (1996) type decompositions (with suitable modifications) have been used since they focus on the covariance between firm size and productivity, and it is more straightforward to link variation in this covariance to economic theory. In addition, structural decompositions of aggregate productivity growth have been developed from economic theory taking into account the shocks and frictions that firms face. These theoretically derived decompositions have been implemented empirically for U.S. manufacturing and show an important contribution of reallocation to productivity growth. The structural decompositions yield a positive contribution of reallocation if inputs are being reallocated on average from low value of marginal product to high value of marginal product plants.
- The HK (2009) and (2017) methodology depends on their identifying distortions and technical efficiency (TFPQ). There are numerous reasons theoretically why the identifying assumptions HK (2009) and (2017) use will not hold. Empirically, multiple

studies using alternative databases and methods suggest that they have not successfully identified either distortions or TFPQ with their indirect methodology. As such, it is difficult to interpret their empirical evidence on the patterns of allocative efficiency in the U.S.

- There is accumulating evidence that dispersion in TFPR and revenue labor productivity is rising substantially in the US particularly in the post 2000 period. Under more general conditions than required by the HK (2009) and (2017) methodology, this evidence is consistent with falling allocative efficiency in the U.S. over this period. What might be driving this? This pattern is consistent with declining business dynamism due to rising reallocation frictions. This is not the inference HK (2017) draw but they don't offer hypotheses or evidence of other factors that can account for the rising revenue productivity dispersion.
- The GHK (2016) and the HK (2017) calibration of creative destruction models help us think about the relationship between models of innovation and the distribution of measured job creation and destruction. However, caution needs to be used in drawing strong inferences based on the work thus far. There are many features of the evidence regarding the contribution of young firms to innovation that are ignored in the current calibration. There are many reasons that young firms face constraints that are ignored in the current calibration. The factors that are ignored suggest HK (2017) understate the contribution of young firms to innovation. Finally, all of their inferences are indirect without any direct empirical evidence of the patterns of within firm productivity growth. This suggests substantial reservations about the finding that most innovation comes from within firm innovations by mature firms.
- In terms of criticisms of the existing arguments and evidence on business dynamism, the biggest limitation of the HK (2017) is a neglect of the potential role of changing reallocation frictions that are dampening both business dynamism and productivity. The type of models emphasized by HK (2017) don't include reallocation frictions in general. This implies that they neglect the possibility that there are increasing reallocation frictions that are impeding the pace of reallocation and lowering productivity in the US. The impact of increasing reallocation frictions can impact productivity growth given that

their impact can take a long time to play out and such increases may exhibit secular trends.

- Returning to the broader theme of the role of reallocation and more generally labor market fluidity for US economic performance, Davis and Haltiwanger (2014) discuss and present evidence that the decline in labor market fluidity has important implications beyond the potential impact on productivity. Their evidence implies that the decline in labor market fluidity has had an adverse impact on labor force participation especially for marginally attached workers. In addition, they discuss the implications of declining labor market fluidity for anemic wage growth. These are consequences potentially related to productivity implications but are important in their own right. In this respect, the focus in HK (2017) only on the productivity implications of declining business dynamism and labor market fluidity is limiting.

Many of these comments have focused on technical details so to conclude let me provide a broader perspective about themes that have been emphasized. First, there are some parts of fluctuations in aggregate productivity that are related to the resiliency of an economy to absorb the constant changes in economic circumstances facing firms and households. These changes include both common and idiosyncratic changes in circumstances. Economies that have greater resiliency are those with the flexibility (broadly defined) to reallocate resources to their most valued use in a timely and cost effective manner. If reallocation frictions increase, productivity will decline, while if they decline, productivity will rise. Policymakers at the local and national level constantly face the challenge of trying to make sure reallocation frictions are not increasing because of market institutions, economic conditions, and the regulatory environment. Sometimes the challenge is that reallocation frictions increase because of economic crises when market institutions like financial markets break down. Since these changes in reallocation frictions can exhibit trends but also because changes in reallocation frictions can take time to have an influence this can result in sustained periods of increased or decreased productivity growth.

At the same time, at the heart of long run productivity growth is innovation. A longstanding open question is the extent to which creative destruction is critical for innovation. We are making progress quantifying this but as the comments above suggest we still don't have either the conceptual framework or especially the data infrastructure to quantify the contribution of creative destruction to innovation.   An example of our limitations is that we don't have a comprehensive data infrastructure that permits tracking within firm productivity growth for the entire U.S. economy.   However, as discussed above, as data on firm growth dynamics is integrated with measures of innovation such as patents and R&D activity the case for creative destruction playing an important role is becoming clearer.

Finally, and importantly, there is an important connection between the reallocation frictions and the contribution of creative destruction for innovation.   If, for example, young firms play a critical role in major innovations, then the barriers they face for entry and post-entry growth are also critical. Thus, an economy that experiences an increase in reallocation frictions can experience a decline in productivity both because resources are being allocated to their highest valued use at a slower pace and because the increase in reallocation frictions impedes the innovation from startups and young firms.   It is from this broad perspective that it is useful to track the pace and nature of business dynamism and labor market fluidity and in turn quantify the mapping between changes in these measures and key economic outcomes like productivity, earnings and economic welfare.

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William R. Kerr, 2013. "Innovation, Reallocation and Growth," NBER Working Paper No. 18993 (April).

Akcigit, Ufuk and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, forthcoming.

Autor, David, John J. Donohue, III, and Stewart J. Schwab, 2006. "The Costs of Wrongful-Discharge Laws," *Review of Economics and Statistics*, 88, no. 2 (May), 211-231.

Autor, David H., William R. Kerr, and Adriana D. Kugler, 2007. "Does Employment Protection Reduce Productivity? Evidence from U.S. States," *Economic Journal*, 117 (June), F189-F217.

Asker, John, Allan Collard-Wexler and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013-1065.

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." In *Brookings Papers on Economic Activity: Microeconomics*, ed. Clifford Winston and Martin Neil Baily. Washington, DC: Brookings Institution Press.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. Cross-Country Differences in Productivity: The Role of Allocation and Selection. *American Economic Review*, 103(1): 305{34, February 2013.

Bils,Mark, Peter J. Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," 2017.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron S. Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2017. "What Drives Differences in Management?" NBER Working Paper No. 23300.

Byrne, David M., John G. Fernald and Marshall B. Reinsdorf. 2016. "Does the United Have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity*, Spring 2016.

Cooper, Russell and John Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies* 73(3); 611-33.

Cooper, Russell, John Haltiwanger, and Jonathan Willis. 2007. "Search Frictions: Matching Aggregate and Establishment Observations." *Journal of Monetary Economics* 54 Supplement 1: 56-78.

Davis, Steven J. and John Haltiwanger. 1999. "Gross Job Flows", *Handbook of Labor Economics*, Volumes 3 and 4, edited by Orley Ashenfelter and David Card, 1999, 2711-2805.

Davis, Steven J. and John Haltiwanger. 2014. “Labor Market Fluidity and Economic Performance.” NBER Working Paper 20479 (published in the 2014 *Federal Reserve Bank of Kansas City Jackson Hole Conference Symposium*).

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. Changing business dynamics and productivity: Shocks vs. responsiveness. 2017.

Elsby, Michael W. L. and Ryan Michaels. 2013. “Marginal Jobs, Heterogeneous Firms, and Unemployment Flows.” *American Economic Journal: Macroeconomics* 5 no. 1:1-48.

Eslava, John Haltiwanger Adriana Kugler Maurice Kugler, Marcela. “Trade and selection: Evidence from manufacturing plants in Colombia.” *Review of Economic Dynamics*, 16(1):135-158, 2013.

Fernald, John. 2014. “Productivity and Potential Output Before, During, and After the Great Recession.” Chap. 1 in *NBER Macroeconomics Annual 2014* edited by Jonathan A. Parker and Michael Woodford, vol. 29, pp. 1-51. Cambridge, MA: MIT Press.

Foster, Lucia, John Haltiwanger, and C.J. Krizan, 2001. “Aggregate Productivity Growth: Lessons from Microeconomic Evidence,” in *New Developments in Productivity Analysis,* edited by Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, 303-363, NBER Studies in Income and Wealth, vol. 63. Chicago: University of Chicago Press.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, 98(1): 394-425.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. “The Slow Growth of New Plants: Learning about Demand.” *Economica*, 83(329): 91-129.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2017. “Macro and Micro Dynamics of Productivity: From Devilish Details to Insights” NBER WP 23666.

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

Haltiwanger, John, Robert Kulick, and Chad Syverson. Misallocation measures: Glowing like the metal on the edge of a knife. 2017.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics*, 124(4): 1403-48.

Goldberg, Penelopi K. and Frank Verboven. 2001. “The Evolution of Price Dispersion in the European Car Market.” *Review of Economic Studies*, 68(4): 811-48.

Gopinath, Gita, Sebnem Kalemli-Ozcan, Loukas Karabarbounis and Carolina Villegas-Sanchez 2017. "Capital Allocation and Productivity in South Europe" *Quarterly Journal of Economics*, (forthcoming).

Hopenhayn, Hugo and Richard Rogerson, 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101, no. 5 (October), 915-938.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403-48.

Kleiner, Morris M. and Alan B. Krueger, 2013. "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market," *Journal of Labor Economics,* 31, no. 2 (Part 2, April), pp. S173-S202.

Melitz, M. J., & Polanec, S. (2015). Dynamic Olley-Pakes productivity decomposition with entry and exit. *The RAND Journal of Economics*, *46*(2), 362–375.

Nakamura, Emi and Dawit Zerom. 2010. "Accounting for Incomplete Pass-Through." *Review of Economic Studies*, 77(3): 1192-1230.

Olley, Steven G. and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263-1297, 1996.

Petrin, Amil and James A. Levinsohn. Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics*, 43(4):705-725, 2012.

Petrin, Amil, T. Kirk White, and Jerome Reiter. The Impact of Plant-level Resource Reallocations and Technical Progress on U.S. Macroeconomic Growth. *Review of Economic Dynamics*, 14(1):3-26, January 2011.

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49(2): 326-65.