Macro and Micro Dynamics of Productivity: From Devilish Details to Insights

G. Jacob Blackwood, Lucia S. Foster, Cheryl A. Grim, John Haltiwanger, Zoltan Wolf*

Abstract

Firm-level revenue-based productivity measures are ubiquitous in studies of firm dynamics and aggregate outcomes. One common measure is increasingly interpreted as reflecting "distortions" since in distortions' absence, equalization of marginal revenue products should yield no dispersion in this measure. Another common, but distinct, measure is the residual of the firm-level revenue function, which reflects "fundamentals". Using micro-level U.S. manufacturing data, we find these alternative measures are highly correlated, exhibit similar dispersion, and have similar relationships with growth and survival. However, the distinction between these alternative measures is critically important for quantitative assessment of the level and decline of allocative efficiency.

Despite broad consensus on the importance of accounting for measured productivity dispersion across firms for understanding aggregate economic performance,[1] there is less consensus about the basics of estimating and interpreting measures of firm-level productivity. Revenue per composite input measures of productivity are the norm in the empirical literature but there are a variety of methods used to compute such measures. The relationship between revenue productivity measures and firm-level fundamentals remains the subject of ongoing debate. We clarify this relationship conceptually and empirically under less strict assumptions about returns to scale, and explore the implications for quantifying allocative efficiency (hereafter AE).

To illustrate the issues of interpretation and the implications for macroeconomics, a good starting point is the set of empirical measures of productivity used by Hsieh and Klenow (2009, 2014) (hereafter HK) to quantify AE. A core insight is that under assumptions of constant-elasticity-substitution (CES) demand and constant-returns-to-scale (CRS) technology, dispersion in revenue per composite input, often referred to as $TFPR$, reflects frictions or distortions impeding the equalization of

*G. Jacob Blackwood: Amherst College, jblackwood@amherst.edu, Lucia S. Foster: U.S. Census Bureau, lucia.s.foster@census.gov, Cheryl A. Grim: U.S. Census Bureau, cheryl.ann.grim@census.gov, John Haltiwanger: University of Maryland, haltiwan@econ.umd.edu, Zoltan Wolf: New Light Technologies, zoltan.wolf@census.gov. Jake Blackwood and John Haltiwanger are also schedule A employees of the U.S. Census Bureau. We thank Jan De Loecker, Ron Jarmin, Kirk White and conference participants at the 2013 Comparative Analysis of Enterprise Data in Atlanta, 2014 Research Data Center Annual Conference, the 2015 NBER CRIW workshop and two anonymous referees for valuable comments. Any remaining errors are our own. Any 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. The statistics reported in this paper have been reviewed and approved by the Census Bureau's Disclosure Review Board (#DRB-B0026-20190205, #CBDRB-FY20-111, #CBDRB-FY20-CES004-025).

[1]See the survey in Syverson (2011).

marginal revenue products. A common approach to calculating the composite input used to compute $TFPR$ is to use cost shares of total costs as input weights.[2] This measure, which we refer to as $TFPR^{cs}$, is readily calculable, and evidence shows that dispersion in $TFPR^{cs}$ varies substantially across countries, industries, and time. Quantifying AE also requires a broad measure of fundamentals reflecting both technical efficiency and demand/product appeal, which HK refer to as $TFPQ$. Under HK assumptions and with estimates of output and demand elasticities, $TFPQ$ is also readily calculable. Despite some apparent conceptual similarities, the properties of $TFPR$ versus $TFPQ$ matter critically for quantifying AE in the HK framework.

How do the HK measures of $TFPR$ and $TFPQ$ relate to measures used in the empirical literature analyzing determinants of firm performance? First, $TFPR^{cs}$ is widely used in the literature as a measure of firm-level performance (i.e., fundamentals) and has been shown to be positively related to firm survival, firm growth, innovation, structured management practices, and exporting behavior. Second, a commonly used alternative is to estimate the relationship between revenue and inputs using regression techniques. To overcome endogeneity issues, control function estimation methods have been developed and become widely used. We call the implied productivity measure that emerges from this estimation procedure $TFPR^{rr}$, where "rr" is short for "revenue function residual." Despite their regular and concurrent use, the distinction between $TFPR^{cs}$ and $TFPR^{rr}$ is often overlooked.

Conceptually, we show that $TFPR^{cs}$ is still reflective of distortions when the CRS assumption is relaxed. Furthermore, we show (log) $TFPR^{rr}$ is proportional to the broadly defined measure of (log) $TFPQ$ by HK, and therefore is also a measure of fundamentals. However, measurement of $TFPQ$, and in turn AE, from revenue and input data requires decomposing revenue elasticities into output and demand components (or more broadly into returns to scale and markup components). We derive a generalized measure of AE under non-constant returns (NCRS) that highlights the importance of decomposing revenue elasticities into their respective components.

Empirically, we begin by computing $TFPR^{cs}$ and $TFPR^{rr}$ using the cost share and the control function approaches, respectively. However, to estimate $TFPQ$ and quantify AE, we also need estimates of the returns to scale and markups. We use two alternative approaches. First, we implement the approach of De Loecker and Warzynski (2012) (hereafter DW) to estimate markups at the industry-level, assuming CRS so that the output elasticities can be estimated using cost shares. Demand elasticities are then computed using cost shares of revenue of variable factors along with the output elasticities.[3] Second, we extend our control function approach by combining it with the relationship between plant-level and industry-level variation as in Klette and Griliches (1996) to decompose revenue elasticities. An advantage of this approach is that it does not impose CRS.

Across these approaches, we find that $TFPR^{cs}$ and $TFPR^{rr}$ exhibit similar dispersion, are highly correlated, and have a similar relationship to key economic outcomes such as firm survival and growth. However, we find that the same underlying data imply very different average sectoral AE (by more than a factor of two) due to differences in elasticity estimates. Moreover, differences in elasticities

[2] Theoretically, the relevant weights are output elasticities divided by returns to scale (RTS).

[3] Unlike De Loecker and Warzynski (2012), we estimate markups at the industry level rather than the plant level.

have dynamic implications, as we find declines in AE for the average industry as well as for aggregate manufacturing from the 1970s to the post-2000 period that range by a factor of more than two.

Two key features reconcile these findings empirically. First, the only difference between $TFPR^{cs}$ and $TFPR^{rr}$ are the weights used to create the composite input. The former uses weights that sum to one while the latter uses revenue elasticities. We find, as does much of the literature, that the sum of the revenue elasticities is below but close to one, on average.[4] Mechanically, this is one of the reasons for similar properties for the two measures. However, there are also conceptual reasons that these two measures are closely related, as we will discuss below.

In contrast, quantifying $TFPQ$ and AE requires decomposing the revenue curvature into the returns to scale and markup components. Although it has been recognized in the literature that increasing curvature generally increases AE, we find it is not only the overall curvature that matters for AE. Increasing the markup and returns to scale in a manner that keeps revenue curvature constant tends to generate lower average AE. More generally, we show that it is the sensitivity of measured dispersion in fundamentals and its correlation with distortions, not dispersion in distortions, which is associated with the sensitivity of AE across elasticity estimates.

The paper is organized as follows. We discuss our methodology and data in Sections I and II. Section III describes the effect of estimation methods on the distribution of elasticity estimates. Section IV describes the implications of the differences in elasticity distributions on productivity dispersion, plant growth and survival, and AE. Section V concludes.

I Methodology

I.A Revenue productivity measures

We specify a Cobb-Douglas production function and a CES demand structure – the core assumptions throughout this section, which are common in the literature.[5] Our formulation is consistent with sectoral output being a CES aggregate of intermediate goods producers given by $Q = \left(\sum_i (\xi_i Q_i)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ where σ is the elasticity of substitution, ξ_i denotes an idiosyncratic demand shifter for plant i, Q_i denotes plant-level quantity and Q industry level quantity. See appendix A.1 for details. Time and sector subscripts are omitted in this section in the notation and equations for expositional convenience. The inverse demand function is given by $P_i = PQ^{1/\sigma}Q_i^{-1/\sigma}\xi_i^{\frac{\sigma-1}{\sigma}}$ for plant i in an industry where P_i denotes plant-level prices and P industry level prices. For expositional ease, we write P_i as $P_i = PQ^{1-\rho}Q_i^{\rho-1}\xi_i^{\rho}$, where ρ=$\frac{\sigma-1}{\sigma}$ with $0 < \rho < 1$.

The plant-level production function is given by $Q_i = \mathcal{A}_i \Pi_j X_{ij}^{\alpha_j}$, where $\mathcal{A}_i$ is technical efficiency, X_{ij} are the plant-level factor inputs (e.g., capital, labor, materials, and energy) and α_j is the elasticity

[4] See, for example, Olley and Pakes (1996), Klette and Griliches (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2006), White, Reiter and Petrin (2011), De Loecker (2011), Gandhi, Navarro and Rivers (2012), and Gopinath et al. (2015).

[5] CES demand is a standard assumption in the productivity literature, see for example, HK, Bartelsman, Haltiwanger and Scarpetta (2013), Foster et al. (2016), and Bils, Klenow and Ruane (2020).

of Q_i with respect to X_{ij}. The log of the revenue function is given by:

$$\log P_i + \log Q_i = \sum_j \beta_j \log X_{ij} + \rho \log \mathcal{A}_i + \rho \log \xi_i + (1-\rho)\log Q + \log P, \tag{1}$$

where the revenue elasticities satisfy $\beta_j{=}\rho\alpha_j$. Various revenue productivity measures have been used in the theoretical and empirical literature. One typical measure is logTFPR, given by (see Foster, Haltiwanger and Syverson (2008)):

$$\log TFPR_i = \log P_i + \log Q_i - \sum_j \alpha_j \log X_{ij} = \log P_i + \log \mathcal{A}_i \tag{2}$$

Equation (2) makes explicit that $TFPR_i$ confounds the effect of output prices and technical efficiency. Decomposing $TFPR_i$ into its price and technical efficiency components is generally not feasible because most micro datasets only contain information about costs and revenues but not plant-level prices. Therefore, the majority of results in the empirical productivity literature are based on revenue productivity measures. An important special case emerges under the assumption that plants minimize total costs and have CRS technology: the share of the jth input expenditure in total costs equals α_{js}. Formally:

$$\log TFPR_i^{cs} = \log P_i + \log Q_i - \sum_j cs_j \log X_{ij} = \log TFPR_i + \sum_j (\alpha_j - cs_j)\log X_{ij}, \tag{3}$$

where cs_j denotes the cost share of the jth input. Note the equivalence between log$TFPR_i$ from equation (2) and log$TFPR_i^{cs}$ does not hold without CRS. Still, log$TFPR_i^{cs}$ is of interest in and of itself, even without CRS, since it is indicative of distortions under certain assumptions, as we demonstrate below. A closely related –but perhaps more subtle– point is that given β_j, implied cost shares can be calculated as $\frac{\beta_j}{\sum_j \beta_j}$ using first order conditions.[6] Since these implied cost shares are determined by the unrestricted β_j estimates, they are not necessarily the same as directly computed cost shares. In what follows, we will compute $TFPR_i^{cs}$ using both approaches.

The revenue productivity measures above are distinct from the revenue function residual which is given by:

$$\log TFPR_i^{rr} = \log P_i + \log Q_i - \sum_j \beta_j \log X_{ij} = \rho \log \mathcal{A}_i + \rho \log \xi_i + (1-\rho)\log Q + \log P, \tag{4}$$

which says that the revenue function residual depends on technical efficiency, idiosyncratic demand shocks and aggregate prices and quantities. In addition, our core assumptions allow $\gamma = \sum_j \alpha_j = \rho^{-1}\sum_j \beta_j$, where γ denotes returns to scale and is not necessarily equal to 1. The implication is that

[6] Another way of stating this point is $\beta_j/\sum_j \beta_j$ yields the output elasticity up to returns to scale. Observe that $\beta_j/\sum_j \beta_j = \alpha_j/\sum_j \alpha_j$.

$TFPR_i^{rr}$ is different from both $TFPR_i$ and its estimate $TFPR_i^{cs}$:

(5) $$\log TFPR_i^{rr} = \log P_i + \log Q_i - \sum_j \beta_j \log X_{ij} \neq \begin{cases} \log TFPR_i \\ \log TFPR_i^{cs}. \end{cases}$$

Without idiosyncratic frictions or distortions, marginal revenue products are equalized across production units and there is no within-industry dispersion in log$TFPR_i^{cs}$. Since this outcome is counterfactual, HK posit the presence of distortions that account for such dispersion. To illustrate this point, we consider the decision problem of firms who maximize static profits with input distortions,[7] which imply:

(6) $$TFPR_i^{cs} \propto \tau_i,$$

where $\tau_i{=}\prod_j (1+\tau_{ij})^{\alpha_j/\gamma}$ denotes a plant-specific weighted geometric average of input distortions and the weights are given by cost shares. Note the proportionality result (6) is obtained equivalently when there are only scale distortions τ_{Q_i}, by using the substitution $\tau_i = (1-\tau_{Q_i})^{-1}$. In contrast, $TFPR_i^{rr}$ is proportional to plant technical efficiency and demand shocks under the same assumptions:[8]

(7) $$TFPR_i^{rr} \propto (\mathcal{A}_i \xi_i)^{\rho}.$$

The key implication for the objective of this paper is that $TFPR_i^{cs}$ is proportional to idiosyncratic distortions while $TFPR_i^{rr}$ is proportional to fundamentals. This conceptual difference, due to using output vs. revenue elasticities, is what motivates the empirical analysis below.

Estimating log$TFPR_i^{rr}$ to measure fundamentals is not novel to this paper. Cooper and Haltiwanger (2006) used revenue function residuals as measures of plant-level fundamentals. The empirical measure of fundamentals used by HK is also tightly linked to this revenue residual approach. To see this, note their empirical measure of $TFPQ$ is equivalent to a composite shock given by $\mathcal{A}_i\xi_i$. That is, their empirical measure of log$TFPQ_i$ is given by: $(\log P_i + \log Q_i)/\rho - \sum_j \alpha_j \log X_{ij}$, which under our core assumptions implies that the indirect measure of log$TFPQ_i = \frac{1}{\rho}\log TFPR_i^{rr}$. Measurement of TFPQ in this indirect fashion is more challenging than measuring log$TFPR_i^{rr}$ since it requires decomposing revenue elasticities into the output elasticity and demand elasticity components.[9]

Under our core assumptions, there need be no systematic relationship between log$TFPR_i^{cs}$ and log$TFPR_i^{rr}$. However, there are many possible distortions that yield a correlation between these measures including size-related distortions, financial constraints, and variable markups that are in-

[7] The profit function in this case is given by $P_iQ_i - \sum_j w_j(1+\tau_{ij}^*)X_{ij}$, where w_j denotes the jth input price.

[8] Here we abstract from industry-level shifters that can be captured by industry-year effects.

[9] In Foster, Haltiwanger and Syverson (2008), $TFPQ$ is defined as the measure of technical efficiency, which is distinct from demand shocks which they also measure. Foster, Haltiwanger and Syverson (2008) and subsequent related literature using P and Q data permit decomposition of the composite shock into its technical efficiency and demand/product appeal components. In this paper, we denote $TFPQ$ as the broader composite shock including both technical efficiency and demand. For current purposes, decomposing the composite shock into its components is not critical. See Foster, Haltiwanger and Syverson (2016) and Eslava and Haltiwanger (2018) for discussion of additional insights that emerge from this decomposition.

creasing in fundamentals. The latter is outside the monopolistic competition CES model but we might view distortions identified in this framework as a reduced form way of capturing all sources of wedges impeding the equalization of marginal revenue products.[10] This broad reduced form view may reflect wedges that are distortions but also may reflect other factors (in the sense they may be outside the control of the social planner). Examples include overhead factors of production that drive a wedge between marginal and average products and adjustment costs that impede responses to shocks to fundamentals.[11]

Our contribution is to explore the systematic relationship between these conceptually distinct productivity measures. We do not formally investigate non-CES demand structures, overhead labor, adjustment costs, or other possible sources of wedges but discuss our findings below in light of the studies that consider these possibilities. Note, when production exhibits non constant returns to scale (NCRS), $\log TFPR_i^{cs}$ is not equal to $\log TFPR_i$. In this case, $\log TFPR_i^{cs}$ will still only reflect any reduced-form distortions while $\log TFPR_i$ will exhibit dispersion even in the absence of such reduced-form distortions. Furthermore, the finding that $\log TFPR_i^{rr}$ is only a function of fundamentals is robust to deviations from CRS. Consequently, we focus on $TFPR_i^{cs}$ and $TFPR_i^{rr}$, since they reflect solely distortions and solely fundamentals, respectively, under the HK framework. $TFPR_i$ itself is reflective of both fundamentals and distortions (aside from the CRS case, where $TFPR_i$ is equivalent to $TFPR_i^{cs}$). While such a measure may be useful in certain contexts, we focus on the measures that clearly distinguish distortions and fundamentals, regardless of returns to scale, which are crucial for measuring AE. Nevertheless, departures from CRS are important for both the estimation of $\log TFPQ_i$ and for measures of AE, which we address below.

I.B Allocative efficiency

Recent literature, beginning with HK and extending through Bils, Klenow and Ruane (2020) (hereafter BKR), builds on the distinction between $TFPQ$ and $TFPR$ using the core assumptions made in the prior section to construct a measure of misallocation which they term allocative efficiency (AE). We revisit these issues since they help highlight the importance of distinguishing between $\log TFPR_{is}^{cs}$ and $\log TFPR_{is}^{rr}$ (since we analyze the effect of aggregation across sectors in this section, variables will be indexed also by s).

Although a key contribution of our paper is to examine AE under NCRS, we first illustrate the framework with firm-level production technology that exhibits CRS, as is standard in the literature.

[10] In the recent literature (see, e.g., Peters (2019), Edmond, Midrigan and Xu (2019), and Baqaee and Farhi (2019)), markup dispersion across producers is the source of idiosyncratic distortions. Variable markups that increase with producer scale is a common approach in this literature.

[11] Another possible source of measured wedges and in turn the correlation between measured $\log TFPR_i^{cs}$ and $\log TFPR_i^{rr}$ is measurement error. For example, measurement error in revenue yields a positive correlation between these measures. Bils, Klenow and Ruane (2020) argue that rising measurement error might account for some of the observed rising correlation between $\log TFPR_i^{cs}$ and $\log TFPQ_{is}$ and the associated declining measured AE in the U.S. manufacturing sector. Decker et al. (2019) find that dispersion in revenue per worker in U.S. manufacturing rises similarly in both survey and administrative data which they argue is not supportive of the rising measurement error hypothesis. We don't attempt to disentangle the sources of wedges in this paper.

We collapse the combined effect of demand shifts and technical efficiency that make up $TFPQ_{is}$ for notational convenience: $A_{is} = \mathcal{A}_{is}\xi_{is}$. That is, as in HK, we think of the empirical measure of fundamentals as a composite of both demand factors and technical efficiency. At the sectoral level, AE is a ratio of sectoral productivity to undistorted sectoral productivity. Sectoral productivity is defined as sectoral output per composite input: $TFPQ_s = Q_s / \prod_j X_{js}^{\alpha_{js}}$. Using CES demand and Cobb-Douglas production with CRS, BKR show that $TFPQ_s$ can be expressed as a power sum of A_{is} weighted by relative distortions:

$$TFPQ_s = \left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s}} \left(\frac{\tau_{is}}{\widetilde{\tau}_s} \right)^{\frac{-\rho_s}{1-\rho_s}} \right)^{\frac{1-\rho_s}{\rho_s}}, \tag{8}$$

where $\widetilde{\tau}_s$ is the harmonic revenue weighted mean of distortions (see Appendix A.2.2). $TFPQ_s$ is maximized when $\tau_{is}{=}\widetilde{\tau}_s$,[12] in which case, following from equation (8), $TFPQ_s$ is given by $A_s^* = \left(\sum_i A_{is}^{\rho_s/(1-\rho_s)} \right)^{(1-\rho_s)/\rho_s}$. AE is defined as the ratio of $TFPQ_s$ to the maximized, counterfactual $TFPQ_s$. Multiplying and dividing by $N_s^{(1-\rho_s)/\rho_s}$, where N_s is the number of plants in the sector, sectoral AE can be expressed as

$$AE_s = \left(\frac{1}{N_s} \sum_i \left(\frac{A_{is}}{\widetilde{A}_s} \right)^{\frac{\rho_s}{1-\rho_s}} \left(\frac{\tau_{is}}{\widetilde{\tau}_s} \right)^{\frac{-\rho_s}{1-\rho_s}} \right)^{\frac{1-\rho_s}{\rho_s}}, \tag{9}$$

where $\widetilde{A}_s = \left(N_s^{-1} \sum_i A_{is}^{\rho_s/(1-\rho_s)} \right)^{(1-\rho_s)/\rho_s}$ is the power mean analogue to A_s^*.

Given our interest in estimation methods that do not impose CRS on plant-level technology, we generalize (9) to be robust to deviations from CRS but otherwise maintain CES demand and Cobb-Douglas production. Our approach, described in detail in Appendix A.2, builds on the appendix of HK –who derive AE_s using a single-input production technology that exhibits decreasing returns to scale– but allows for multiple inputs and also NCRS. This generalization is useful as it helps us draw out the implications of the alternative estimation approaches for AE. Under these less restrictive assumptions, $TFPQ_s$ is given by:[13]

$$TFPQ_s = \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s\gamma_s}} \left(\frac{\tau_{is}}{\widetilde{\tau}_s} \right)^{\frac{-\rho_s\gamma_s}{1-\rho_s\gamma_s}} \right)^{\frac{1-\rho_s\gamma_s}{\rho_s}}}{\left(\prod_j X_{js}^{\alpha_j/\gamma_s} \right)^{1-\gamma_s}}. \tag{10}$$

Let X_{js}^* denote aggregate input j corresponding to $\max\{TFPQ_s\}$, the case where distortions are

[12] More details are available in Appendix A.2.4, which also shows the sufficient condition is satisfied only if $\rho_s\gamma_s{<}1$. This is an intuitive restriction, as an equilibrium with increasing returns in the revenue function would imply one firm taking over the market. However, not all estimation methods restrict the parameter space to ensure this is the case.

[13] Equation (10) is a generalization of the equation on page 1445 of the appendix in HK.

equalized across plants. Dividing and multiplying by N_s appropriately, AE_s can then be obtained as:

$$(11) \qquad AE_s = \underbrace{\left(\frac{1}{N_s}\sum_i \left(\frac{A_{is}}{\widetilde{A}_s}\right)^{\frac{\rho_s}{1-\rho_s\gamma_s}} \left(\frac{\tau_{is}}{\widetilde{\tau}_s}\right)^{\frac{-\rho_s\gamma_s}{1-\rho_s\gamma_s}}\right)^{\frac{1-\rho_s\gamma_s}{\rho_s}}}_{AE_s^{COV}} \underbrace{\left(\frac{\prod_j X_{js}^{*\alpha_{js}}}{\prod_j X_{js}^{\alpha_{js}}}\right)^{\frac{1-\gamma_s}{\gamma_s}}}_{\text{Sectoral Intermediate Term}} .$$

Equation (11) is a generalization of (9) that fully accounts for the effect of NCRS production technology.[14] The first term – labeled as AE_s^{COV} in order to emphasize that it resembles a covariance term– shows the effect of NCRS on the within-industry component of AE. The second term in (11) captures the effect of NCRS via sectoral inputs. Importantly, this term equals 1 when all production factor supplies are exogenous, implying that only AE_s^{COV} is relevant in this case.

It is instructive to highlight the role of ρ_s, γ_s, and α_{js} for sectoral AE. Equation (11) shows that these parameters affect AE_s^{COV} via the exponents. In addition, ρ_s and γ_s affect relative technical efficiencies and distortions since both A_{is} and $TFPR_{is}^{cs}$ are constructed using an input index as the denominator where the input index depends directly on α_{js} estimates. The implication is that the joint distribution of these variables is a key determinant of AE. In order to formalize this result, we express AE_s^{COV} as a function of the covariance between transformations of τ_{is} and A_{is}:[15]

$$(12) \qquad \log AE_s^{COV} = \gamma_s \log\left(\frac{\widetilde{\tau}_s}{\overline{\tau}_s}\right) + \frac{1-\rho_s\gamma_s}{\rho_s}\log\left[\operatorname{cov}\left(\left(\frac{A_{is}}{\widetilde{A}_s}\right)^{\frac{\rho_s}{1-\rho_s\gamma_s}}, \left(\frac{\tau_{is}}{\overline{\tau}_s}\right)^{\frac{-\rho_s\gamma_s}{1-\rho_s\gamma_s}}\right) + 1\right].$$

Equation (12) reveals AE_s^{COV} depends on sectoral distortions (term 1), and a function of the covariance between exponentiated relative technical efficiencies and distortions (term 2). By definition, the covariance (term 2) depends on the dispersion of these two variables and the correlation between them. This relationship creates a useful link between AE_s and the properties of the within-industry productivity distribution. For example, if distortions are positively correlated with fundamentals, as is increasingly assumed in the literature, then this component of AE is decreasing in dispersion of either A_{is} or τ_{is}. This can be seen by noting that the dispersion of $\left(A_{is}/\widetilde{A}_s\right)^{\frac{\rho_s}{1-\rho_s\gamma_s}}$ and $(\tau_{is}/\overline{\tau}_s)^{\frac{-\rho_s\gamma_s}{1-\rho_s\gamma_s}}$ in equation (12) are both increasing in the dispersion of A_{is} and τ_{is} but the negative exponent of $\tau_{is}/\overline{\tau}_s$ implies the positive correlation between A_{is} and τ_{is} translates into negative correlation between the transformed variables. Furthermore, changes in the correlation reinforce these patterns.[16]

We use equation (12) in our empirical analysis below to provide guidance about the sensitivity of the measures of AE to the estimates of ρ_s, γ_s, and α_{js}. The second term of equation (12) highlights

[14] Note, (11) simplifies to (9) under CRS ($\gamma_s = 1$).

[15] Equation (12) can be obtained by multiplying and dividing AE_s^{COV} by $\overline{\tau}_s = \left(\frac{1}{N_s}\sum_i \tau_{is}^{\frac{\rho_s\gamma_s}{\rho_s\gamma_s-1}}\right)^{\frac{\rho_s\gamma_s-1}{\rho_s\gamma_s}}$ and rearranging the resulting expression.

[16] Hopenhayn (2014b,a) argues that the conventional wisdom that an increase in the correlation between fundamentals and distortions decreases AE is not necessarily correct. However, he notes that calibration by Restuccia and Rogerson (2008) supports this conventional wisdom and also that this wisdom holds as long as the initial dispersion of distortions is sufficiently large. The second term of equation (12) unambiguously declines with an increase in the correlation between fundamentals and distortions but the first term is also potentially important. We show empirically below that the second term dominates, supporting the conventional wisdom.

the complex relationship between curvature parameters and AE. Analytic derivatives of this second term (available upon request) with respect to ρ_s and γ_s deliver an ambiguous sign that depends on the sign of the covariance between fundamentals and distortions. Critically, ρ_s and γ_s do not enter AE symmetrically, so the influence of the two parameters cannot be summarized by revenue curvature. We show in Appendix A.2.7 that while the plant-level responses to productivity shocks depend on revenue curvature, ρ_s also impacts sectoral responses through the CES aggregator, generating an asymmetry in the influence of the two parameters. Further complicating matters, the empirically measured distributions of fundamentals, including the correlation between fundamentals and distortions, depends asymmetrically on ρ_s and γ_s. Mechanically, for given revenue and input expenditure data, the variance of the numerator of measured $TFPQ_{is}$ is decreasing in ρ_s with a multiplicative factor that depends on the dispersion of measured revenue while the variance of the measured denominator is increasing in γ_s with a multiplicative factor that depends on the dispersion of measured inputs. This in turn implies the covariance of measured $TFPQ_{is}$ with measured $TFPR^{cs}_{is}$ varies with ρ_s and γ_s in an asymmetric manner.[17] The dependence of the distribution of fundamentals provides a concrete channel through which we can characterize the influence of parameter estimation on inferred AE. In our empirical analysis, we show how this sensitivity of the key moments to the estimates of ρ_s and γ_s is quantitatively important for inferences about AE. Since equation (12) also shows these estimates matter for AE for a given distribution of distortions and fundamentals, we use this decomposition to quantify how each of the terms varies with these estimates empirically.

It is also instructive to highlight the connection between $\log AE^{COV}_s$, $\log TFPR^{cs}_{is}$, and $\log TFPR^{rr}_{is}$. Critical here is that $\log A_{is} = \frac{1}{\rho_s} \log TFPR^{rr}_{is}$. Since $\tau_{is} \propto TFPR^{cs}_{is}$ in practice, the empirical equivalent of equation (12) can be re-written as:[18]

$$(13) \qquad \log AE^{COV}_s = \gamma_s \log \frac{\widetilde{TFPR}^{cs}_s}{\overline{TFPR}^{cs}_s} + \frac{1-\rho_s\gamma_s}{\rho_s} \ln \left[\operatorname{cov} \left(\left(\frac{TFPR^{rr}_{is}}{\widetilde{TFPR}^{rr}_s} \right)^{\frac{1}{1-\rho_s\gamma_s}}, \left(\frac{TFPR^{cs}_{is}}{\overline{TFPR}^{cs}_s} \right)^{\frac{-\rho_s\gamma_s}{1-\rho_s\gamma_s}} \right) + 1 \right].$$

In other words, measured AE^{COV}_s is a function of the two distinct revenue productivity measures derived above. We return to this relationship in our empirical analysis below.

For the most part, we focus on AE at the sectoral level. However, it is helpful to explore various methods aggregating AE across sectors. Following the literature (e.g, HK and BKR), we treat supply of aggregate primary factors as fixed and assume a Cobb-Douglas CRS aggregator for output across sectors into a final good. In addition, we assume a representative perfectly competitive firm that produces this final output. It can be shown that these two assumptions imply that primary sectoral inputs are constant so long as average industry distortions are unchanged. Thus their contribution to AE drops out of (11).

Alternative approaches can be used with respect to treatment of intermediate inputs. One approach is to treat intermediates as inputs that are in fixed supply external to the sectors under consideration. This would hold for raw materials and energy. For a subsample of manufacturing,

[17] In contrast, measured $TFPR^{cs}_{is}$ does not directly depend on ρ_s and γ_s.

[18] In this expression we use the same type of aggregation as in equation (12).

this could be considered to hold for the intermediate inputs produced independently of the sectors under consideration.[19] Sectoral AE simplifies to AE_s^{COV} in this case and overall AE is given by:

$$AE = \prod_{s}^{S} AE_s^{\theta_s} = \prod_{s}^{S} \left(AE_s^{COV}\right)^{\theta_s}, \tag{14}$$

where θ_s denotes the revenue share of industry s. A second approach, labeled roundabout production, endogenizes intermediate input production recognizing that goods are used to produce goods. BKR implement this in a simplified manner where all intermediates are taken from aggregate output of the sectors under consideration.[20] In this case, while inputs do not completely drop out, we can still express AE in industry s as a function of AE_s^{COV}:[21]

$$AE_s = \left(AE_s^{COV}\right)^{\frac{\sum_{k=1}^{S}\theta_k\left(1-\frac{\alpha_{Mk}}{\gamma_k}\right)}{\sum_{k=1}^{S}\theta_k\left(1-\frac{\alpha_{Mk}}{\gamma_k}\right)+\theta_s\frac{\alpha_{Ms}}{\gamma_s}(1-\gamma_s)}}, \tag{15}$$

where α_{Ms}/γ_s denotes the cost share of intermediate inputs in industry s and AE_s^{COV} is defined in equation (11).[22] Aggregation across sectors implies the following expression:

$$AE = \prod_{s=1}^{S} AE_s^{\frac{\theta_s}{\sum_s \theta_s\left(1-\frac{\alpha_{Ms}}{\gamma_s}\right)}} = \prod_{s=1}^{S} \left(AE_s^{COV}\right)^{\frac{\theta_s}{\sum_{k=1}^{S}\theta_k\left(1-\frac{\alpha_{Mk}}{\gamma_k}\right)+\theta_s\frac{\alpha_{Ms}}{\gamma_s}(1-\gamma_s)}}. \tag{16}$$

These two aggregate concepts are relevant for the empirical analysis because they act as a form of upper and lower bounds on AE, since roundabout production amplifies the effects of distortions (with or without CRS). Equation (14) overstates true AE if some industries are characterized by roundabout production. On the other hand, equation (16) understates true AE if there are industries where intermediate inputs are in fixed supply. Given these measures provide bounds, we show results using both approaches.

I.C Estimation methods

We begin by reviewing the cost-share-based method for estimating output elasticities that we use in our estimates of $TFPR_{is}^{cs}$.[23] Cost-share-based methods (CS) exploit first order conditions from the firm's cost-minimization problem. This framework implies that under CRS, the share of input expenditures in total costs identify output elasticities even without data on prices and quantities.

[19] HK finesse this point by using value-added production so they only have primary inputs.

[20] We adopt the aggregation of BKR to allow for endogenous intermediates for this purpose but note that more complex input-output linkages are of interest in this context (see, e.g., Baqaee and Farhi (2019)). Although we treat the model with fixed inputs and roundabout production as bounds on the influence of intermediates on AE, models with more complex input-output linkages could yield a measure of AE outside this range.

[21] See appendix A.2.6 for details on equations (14) and (15).

[22] Under CRS, the outer exponent simplifies to 1, and therefore (15) collapses to (9). When returns to scale are increasing, the exponent is greater than one and AE_s^{COV} is smaller relative to when inputs are exogenous, because the influence of intermediates serves to amplify the effect of the inner term. On the other hand, decreasing returns increases measured AE_s^{COV}.

[23] We follow the standard practice for both $TFPR_{is}^{cs}$ and $TFPR_{is}^{rr}$ by deflating plant-level revenue with the industry-level deflator. Accordingly these are measures of real revenue per composite input using industry-level deflators. This implies that we have moved the industry-level deflator to the LHS of equation (1).

This is a useful property because it makes CS robust to alternative demand structures and also imperfect competition in output markets. A potential caveat is that an estimate of, or an assumption about, returns to scale is necessary. Since the requirement that the first order conditions of cost minimization hold for all businesses in every time period may be considered too restrictive, most studies (see Syverson (2011)) average them across plants in an industry and/or over time. This approach is equivalent to assuming elasticities are homogenous within an industry or over time, a restriction that alternative estimation methods also typically require.

In order to estimate the revenue elasticities and $TFPR_{is}^{rr}$, we use a control function approach. The method was developed by Olley and Pakes (1996) (hereafter OP) to address the endogeneity of unobserved productivity and inputs.[24] Our version combines constrained optimization and OP, which is why we label our method as "OPH" where "H" is short for "hybrid". The main difference relative to OP is that first order conditions of profit maximization are used to identify the revenue elasticities of variable inputs. Specifically, we estimate industry-specific revenue elasticities for variable factors (intermediate inputs and energy) as the mean of the plant-level revenue share of input expenditures. Next the contribution of variable inputs is subtracted from revenue variation using these revenue shares, and the remaining variation is used to determine the elasticities of capital and labor using the control function approach. Since quasi-fixed input decisions have dynamic consequences, their elasticities are determined in an additional step using the residual revenue variation that is left after removing the contribution of variable inputs. Note that in order to account for the possible endogeneity of productivity and labor, we treat labor as quasi-fixed, which is an additional deviation from the original OP.[25] The identifying assumptions underlying control function estimators have been criticized.[26] Our approach is robust to these criticisms because we use plant-level investment as a proxy, and variable input elasticities are not determined by projecting revenue variation on potentially endogenous inputs.[27]

Without detailed data on output price variation or further assumptions, control function approaches in general yield revenue elasticities and not output elasticities.[28] This distinction – either stated as a bias due to unobserved price variation or neglected altogether in the applied literature – is

[24]The procedure, now commonly used in the applied literature in various forms, is based on the assumption that investment is monotonically increasing in the composite revenue shock, which is treated as the only unobserved state variable. Under this assumption, the control function is invertible and the estimated plant-level composite revenue shock can be used to control for endogenous productivity differences. An earlier working paper version of this paper Foster et al. (2017) explored other control function/proxy methods and obtained very similar results.

[25]We use a third-order polynomial in state variables and the proxy to control for the unobserved residual. Under OPH, the residual includes the composite revenue shock and industry-level real output as in (1). Recall we control for the industry-level price deflator by using real revenue. We treat this residual as composite unobserved state variable and we follow Olley and Pakes (1996) for the final estimation step using nonlinear least squares using their proxy and selection correction terms.

[26]See, for example, Levinsohn and Petrin (2003), Wooldridge (2009), Gandhi, Navarro and Rivers (2012), Ackerberg, Caves and Frazer (2015), and the NBER working paper version of this paper for more discussion on these matters. Our procedure can be interpreted as a non-parametric alternative to Gandhi, Navarro and Rivers (2012).

[27]See Haltiwanger and Wolf (2018) for more details. A version for CES technology is discussed in Dinlersoz and Wolf (2018).

[28]An exception is Eslava and Haltiwanger (2018) which uses plant-level price and quantity data to jointly estimate the production and demand functions using GMM procedures motivated by the control function approach.

critical for our purposes since it implies that the residual $TFPR_{is}^{rr}$ is a combination of fundamentals reflecting both demand and production characteristics. In order to be able to make inference about output elasticities and gain further insights, we consider two separate estimates of ρ_s. The first approach follows the method described in De Loecker and Warzynski (2012). Specifically, the first order condition for a variable factor, such as intermediate inputs, yields that the markup ($1/\rho_s$) is equal to the ratio of the output elasticity to the cost share of revenue of the variable factor. The challenge then is how to estimate the output elasticity. We use the cost share of total costs for materials at the industry level to obtain the output elasticity associated with our CS methodology. This approach, which we denote by DW, yields estimates of markups (and thus ρ_s) that vary across industries.[29] DW is fully consistent with our CS methodology and potentially consistent with OPH.

A second approach uses the method described in Klette and Griliches (1996) combined with our control function approach. This specification is labeled "OPHD" where "D" is short for "Demand". In addition to estimating the output elasticities for capital and labor using the control function approach we also estimate ρ_s using an auxiliary regressor of industry-level log real output. OPHD has the advantage of estimating revenue elasticities and demand elasticity in an internally consistent manner without imposing restrictions such as CRS on output elasticities.[30]

II Data

II.A Source data

Our industry-level data, including deflators, capital rental prices, and depreciation rates, are taken from the NBER-CES Manufacturing Industry Database, the Bureau of Labor Statistics, and the Bureau of Economic Analysis. We use plant-level information from the Annual Survey of Manufactures (ASM), Census of Manufactures (CM), and the Longitudinal Business Database (LBD).[31]

We use the ASM and CM to construct plant-level measures of inputs and real revenue. Real revenue is measured as the deflated total value of shipments, corrected for the change in finished goods and work-in-process inventories. Total hours worked is constructed as the product of production worker hours and the ratio of the total wage bill to production worker wages. Our intermediate input variable is given by the sum of three items: cost of parts, contracted work, and goods resold.

[29]Estimating output elasticities in this manner requires strong assumptions including CRS. De Loecker and Warzynski (2012) use a control function approach that does not impose CRS. However, since they do not observe plant-level prices their estimates of output elasticities are more appropriately interpreted as revenue elasticities. De Loecker and Warzynski (2012) also estimate plant-level markups using the first-order condition for variable factors.

[30]Under OPHD, we include industry-level real output as a regressor consistent with equation (1) so the unobserved residual in this specification only reflects the composite revenue shock. Industry-level real output is treated like capital and labor in the Olley and Pakes (1996) selection and proxy corrections. OPHD can be interpreted as a refinement of OPH that is a more robust procedure for estimating revenue elasticities (e.g., the selection and proxy corrections are more general including industry-level real output). The revenue elasticities for variable inputs are identical for OPH and OPHD and the estimates for the quasi-fixed factors are very similar under OPH and OPHD. Thus, the difference in methodology between OPH and OPHD yields little differences in revenue elasticities. We discuss these issues further below.

[31]More information on the construction of the database is available in the online data documentation. An earlier version of the NBER-CES database is documented in Bartelsman and Gray (1996).

The energy input consists of deflated electricity and fuel costs. We create plant-level capital stock measures using a version of the Perpetual Inventory Method, which calculates current capital as a sum of the depreciated stock and current investment.

The LBD serves two purposes in our analysis. First, high-quality longitudinal identifiers help us determine the accurate time of plant exit which is needed to estimate the relationship between productivity, growth, and exit. Second, the LBD acts as a universe file; we use employment and plant age data from the LBD to construct inverse propensity score weights that control for selection to the ASM. More details about the data can be found in the working paper version of this study and Appendix A in Foster, Grim and Haltiwanger (2016) (FGH, hereafter). These descriptions include how cost shares of inputs are measured.

II.B Analysis samples

Our initial sample includes approximately 3.5 million plant-year observations between 1972 and 2010. Although this is a large dataset, we restrict the sample in the empirical analysis because it needs to fulfill two conflicting requirements. First, industries should be defined narrowly enough that we can plausibly assume elasticities are constant across plants within an industry. We regard this as an important requirement since dispersion in measured revenue productivity measures across plants in the same industry may be due unmeasured heterogeneity in elasticities – this issue is especially relevant if broad sectoral definitions are used. To fulfill this requirement, we choose a narrow 4-digit SIC grid (corresponding to a 6-digit NAICS grid). Second, the number of plant-year observations within each industry should be large enough that elasticities can be estimated by OPH and OPHD. Large samples are necessary for two reasons. First OPH and OPHD use high-order polynomials making estimates sensitive to small samples, see section I.C. Second, the selection correction is based on internally estimated exit probabilities which require a sufficiently large number of exit observations. Empirical studies often use a 2 or 3-digit industry grid in order to be able to generate sensible elasticities for all industries. We wanted to avoid this so that we could compare and contrast OPH, OPHD, and CS using detailed industries.

Changes in industry classification systems over time create complications because they entail spurious breaks in plant-level time series and a drop in sample size. We address this by selecting 4-digit SIC industries which were either not affected by classification changes or mapped one-to-one into another SIC category (in 1987) or NAICS category (in 1997). There are 292 such industries of which we selected the first 50 based on the number of plant-year observations.[32]

While our primary analytic sample is the 50 most populous industries, we also analyze a total manufacturing sample covering all industries from 1972-2010. This analysis is restricted to the CS method for reasons discussed above. The total manufacturing sample is of particular interest and importance for sensitivity analysis of aggregate AE.

[32] According to the NBER-CES Manufacturing Industry Database, this industry set accounts for about 36% percent of total Manufacturing value added between 1972 and 2010 with very little annual volatility.

III Elasticity distributions

We start by estimating output and revenue elasticities for each of our four inputs: capital, labor, energy, and materials. Under both OPH and OPHD, the revenue elasticities of variable inputs, β_{Es} and β_{Ms}, are obtained as the revenue share of respective expenditures. Therefore, β_{Es} and β_{Ms} are the same for OPH and OPHD, see Figures 1(a)-1(b). Output elasticities α_{Es} and α_{Ms} are calculated as the share of input expenditures in total costs under CS. Note, β_{Es} and β_{Ms} tend to be lower than α_{Es} and α_{Ms}. This is to be expected since under CRS technology and CES demand the revenue elasticities should be lower than output elasticities.[33]

Figures 1(c)-1(d) plot the estimated elasticity of output (CS) and revenue (OPH, OPHD) with respect to capital and labor (β_{Ks} and β_{Ls}), distributed across industries. There are non-trivial differences in both the location and the shape of the distributions. Most notably, the CS-based capital and labor elasticities tend to be smaller than regression-based estimates. At first glance, this is contrary to expectations since under CRS technology and CES demand the revenue elasticities should be lower than output elasticities. However, the capital and labor elasticities under OPH and OPHD are determined without using information either on total costs or the relationship between revenue and cost shares, and there is no restriction forcing the above relationship to hold. Observe as well that capital and labor revenue elasticities under OPH and OPHD are very similar. In light of this similarity, it is useful to highlight that OPHD is interesting because it permits decomposing revenue elasticities into factor elasticities and the demand elasticity in an internally consistent manner.

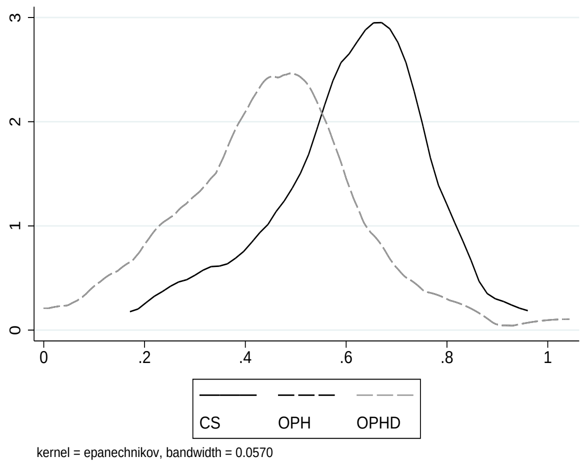
Figures 2(a)-2(c) summarize the overall implications of the differences between OPH and OPHD. Figure 2(a) confirms that the similar revenue elasticities under OPH and OPHD (figure 1) imply consequently similar revenue curvature distributions. The sample averages of $\sum_j \widehat{\beta}_{js}$ under OPH and OPHD, respectively, are 0.94 and 0.95. The standard deviation is 0.24 for both.[34] These patterns are broadly consistent with those in the literature: estimated revenue curvature in the average industry was found to be close to 1 in many other papers.[35] This common finding is important for the properties of $\log TFPR^{rr}_{is}$, as we show below.

We decompose the curvature of the revenue function into demand elasticities (ρ_s) and returns to scale (γ_s) using two ρ_s distributions, see section I.C. The first ρ_s distribution we examine, based on De Loecker and Warzynski (2012), and denoted by DW, uses industry-level data from the NBER-CES Manufacturing Industry Database to calibrate ρ_s under the assumption that γ_s=1. Since this restriction is removed under OPH, we abstract from it for the sake of this exercise and condition on ρ_s as if it were exogenous in order to determine the γ_s implied by $\sum_j \beta_{js}$. In the second ρ_s distribution, denoted by OPHD, ρ_s is estimated jointly with β_{js} using micro data following Klette

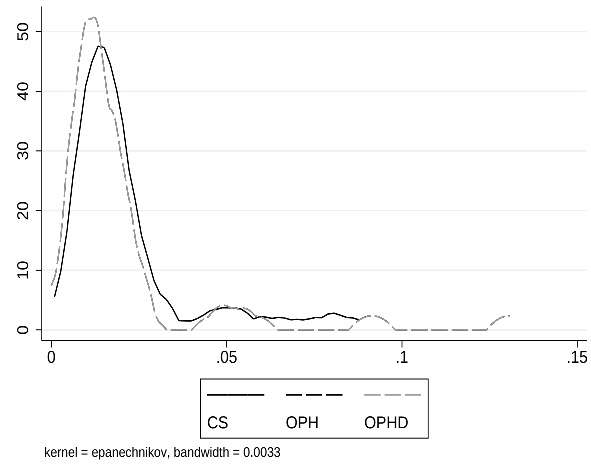
[33] The revenue and total cost based cost shares of intermediate inputs and energy have correlations of 0.67 and 0.88, respectively.

[34] The Kolmogorov-Smirnov test indicates that the two distributions are not significantly different.

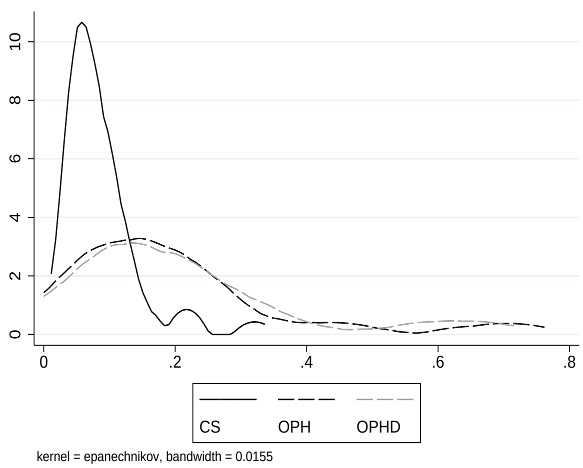
[35] See, for example, Olley and Pakes (1996), Klette and Griliches (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2006), White, Reiter and Petrin (2011), De Loecker (2011), Gandhi, Navarro and Rivers (2012), and Gopinath et al. (2015).



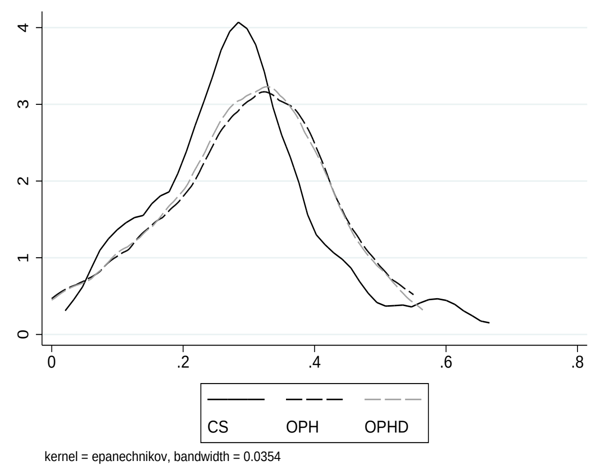
(a) $\widehat{\beta}_{Ms}$ and $\widehat{\alpha}_{Ms}$



(b) $\widehat{\beta}_{Es}$ and $\widehat{\alpha}_{Es}$



(c) $\widehat{\beta}_{Ks}$ and $\widehat{\alpha}_{Ks}$

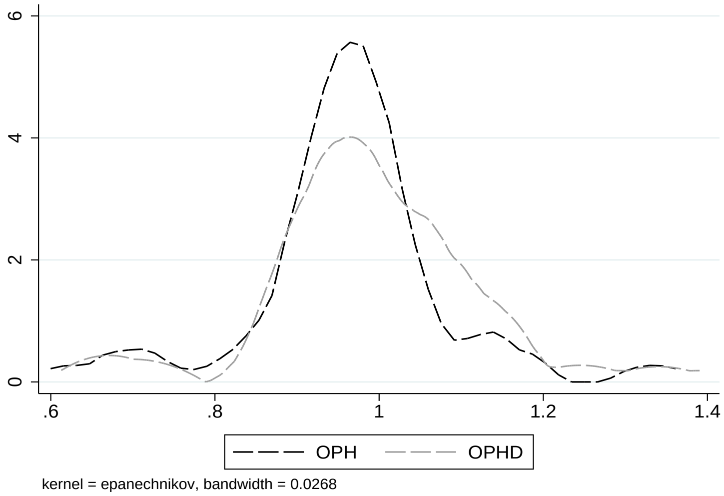


(d) $\widehat{\beta}_{Ls}$ and $\widehat{\alpha}_{Ls}$

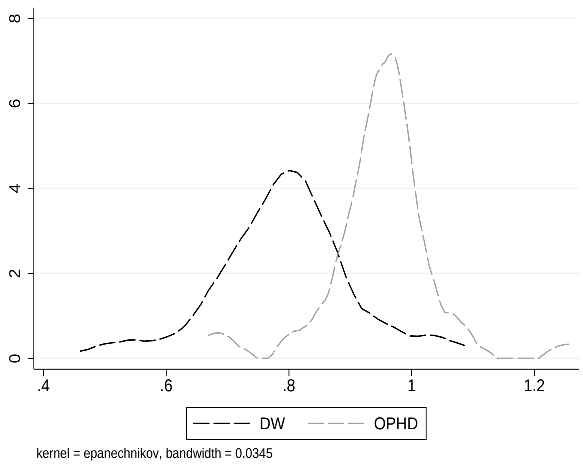
Figure 1: Cross-industry distributions of output (α_{js}) and revenue elasticities (β_{js})
Note: $\widehat{\beta}_{Ms}$ and $\widehat{\beta}_{Es}$ are the same for OPH and OPHD, by construction. CS yields $\widehat{\alpha}_{js}$.

and Griliches (1996). Figure 2(b) plots the two ρ_s distributions. Although the heterogeneity is similar, DW-based ρ_s-s are significantly lower than OPHD-based ones.[36] The disparate ρ_s estimates have important implications for γ_s, shown in figure 2(c). Under OPHD, mean γ_s is close to one (1.01), while it is significantly larger if ρ_s-s are calibrated (with mean 1.21, denoted as OPH in this figure). Both γ_s distributions exhibit considerable variation across industries with some industries exhibiting decreasing returns to scale and others showing increasing returns to scale. In particular, using a delta-method-based Wald test, we reject the null of CRS in roughly two-thirds of the industries under OPH and three-quarters of the industries under OPHD: 60% of industries show increasing and 4% show decreasing returns to scale under OPH; the corresponding fractions are 49% and 26% under OPHD. In other words, both OPH- and OPHD-based revenue elasticities imply increasing returns to scale

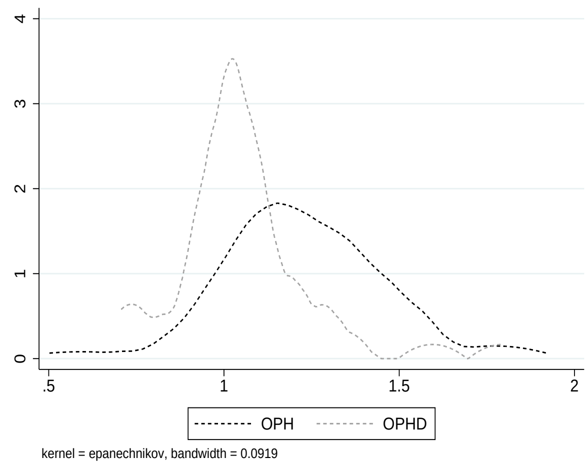
[36]The standard deviations are 0.11 and 0.09, the means are 0.8 and 0.94. The Kolmogorov-Smirnov test confirms that the difference between the two distributions is statistically significant.



(a) Revenue function curvature: $\widehat{\rho_s \gamma_s} = \sum_j \beta_{js}$



(b) Demand elasticity estimates (ρ_s) as in De Loecker and Warzynski (2012) and Klette and Griliches (1996)



(c) Implied returns to scale (γ_s)

Figure 2: Revenue curvature and its components

in a nontrivial number of industries. Interpreted as a specification-check, the results of these tests present the tension between microdata-based elasticity estimates and the CRS-assumption necessary to derive ρ_s.

Taking stock, the control function approach yields substantial heterogeneity in revenue function curvature across industries with the average just below one. This heterogeneity may be amplified into a wide range of returns to scale estimates, depending on the underlying demand elasticities. We consider the implications of these findings for the firm dynamics literature that requires calibrating revenue curvature. To help preserve well-behaved optimization problems, it is typical to assume that $\rho_s \gamma_s \leq 1$. Our findings are not inconsistent with this assumption but indicate that estimated curvature and its sources may exhibit significant variation, depending on the estimation method. If the primary source of curvature is the markup and it is substantial, i.e. in the 25 percent range (ρ_s=0.8), then the control function approach implies that the returns to scale for production are well above one. This inference is not limited to our findings since the revenue curvature is commonly found to be close to or just below one. In short, the control function estimates imply that either markups are small or returns to scale of production are above one. Widespread recognition of this implication has been limited with estimates of β_{js} frequently interpreted as estimates of α_{js} in the literature.[37]

IV Implications of the differences in elasticity distributions

We now turn to exploring the effects of the differences in elasticity distributions on the basic properties of revenue productivity measures and AE. In particular, we explore the effect of these differences in terms of productivity dispersion, productivity correlations, and the relationship between productivity, growth, and survival. We also investigate the sensitivity of AE to these different approaches to estimating factor and revenue elasticities.

IV.A Productivity dispersion and correlations

Does it matter whether one uses output or revenue elasticities to compute the composite input and, in turn, does it matter how one estimates the output and revenue elasticities? In spite of the large differences in elasticity estimates presented above, our results suggest that, at least on average, dispersion in revenue productivity is broadly similar across methods. The interquartile range, shown in the second column of Table 1, indicates that the average productivity difference between establishments at the 75th and 25th percentiles in the average industry varies between 0.28 and 0.35 across the methods considered for the purposes of this paper.[38] This narrow range amounts to a 31-42% productivity difference, indicating substantial within-industry dispersion in revenue productivity. When measuring dispersion using the standard deviation, the results are qualitatively the same (see the third column of Table 1).

Table 1: Productivity dispersion implied by different methods

	Sample size (in 1000)	Interquartile range	Standard deviation
CS	589	0.28	0.31
OPH	563	0.35	0.35
OPHD	563	0.35	0.36

All statistics are based on deviations of plant-level log-productivity from industry- and time-specific means and are calculated from a weighted distribution where the weights are based on the number of plant-year observations in an industry. The top and bottom 1% of the distributions are trimmed. Source: authors' calculations.

We next investigate whether the choice of estimation method has consequences for the productivity rank of establishments. The Pearson and Spearman correlations in Table 2 indicate that the

[37] There are, of course, exceptions. De Loecker (2011) was one of the first to be careful about the interpretation of the control function estimates as reflecting revenue rather than output elasticities.

[38] Approximate 95% confidence intervals, constructed using bootstrapped standard errors of the interquartile range, not shown here, indicate that dispersion measures under the control function approach are higher than under CS but they are not significantly different from each other.

association between log$TFPR_{is}^{rr}$ and log$TFPR_{is}^{cs}$ is generally weaker than between log$TFPR_{is}^{rr}$ implied by different control function approaches themselves, but all correlations are higher than 0.7. These findings suggest that $TFPR_{is}^{cs}$ and $TFPR_{is}^{rr}$ exhibit similar dispersion and are strongly correlated. The dispersion in $TFPR_{is}^{rr}$ tends to be somewhat larger than $TFPR_{is}^{cs}$. Under the assumption of isoelastic (CES) demand, $TFPR_{is}^{rr}$ is a measure of fundamentals. Consistent with the findings of the recent literature, our results imply that whether or not $TFPR_{is}^{cs}$ is an appropriate measure of distortions, it is positively correlated with, and similarly dispersed as, fundamentals.[39] Mechanically, part of what is driving this close correspondence between $TFPR_{is}^{cs}$ and $TFPR_{is}^{rr}$ is that they (at least on average) both generate the composite input using weights that sum close to one. Conceptually, as we have discussed above, there are number of reasons for these alternative measures to be positively correlated.

Table 2: Correlations among within-industry productivity distributions

	CS	OPH	OPHD	CS	OPH	OPHD
	Pearson			Spearman		
CS	1			1		
OPH	0.73	1		0.74	1	
OPHD	0.71	0.88	1	0.71	0.90	1

The top and bottom 1% of the distributions are trimmed. Source: authors' calculations.

IV.B Growth and survival

In this sub-section, we explore whether one of the most important predictions from standard models of firm dynamics is robust to the differences in elasticity distributions. For this purpose, we are motivated by standard models in which firms face adjustment frictions on both the entry/exit and intensive margins.[40] In such a model where employment is the single production factor subject to adjustment frictions, incumbent firms have two key state variables each period: the prior period level of employment and the realization of productivity in the period. The standard prediction from this model is that, conditional on prior period level of employment, firms with sufficiently low draws from the productivity distribution exit and firms with higher realizations of productivity grow. Syverson (2011) highlights that the positive relationship between productivity, growth, and survival is a robust finding in the literature. If the productivity measure is $TFPR_{is}^{rr}$, this should not surprise us because it reflects fundamentals. However, it is less clear the prediction should hold for $TFPR_{is}^{cs}$, given our

[39] These findings are also consistent with studies that use price and quantity data to compute direct measures of technical efficiency and demand shocks. Foster, Haltiwanger and Syverson (2008), Foster, Haltiwanger and Syverson (2016), Haltiwanger, Kulick and Syverson (2018), and Eslava et al. (2013) provide evidence that $TFPR_{is}^{cs}$ is highly correlated with direct measures of technical efficiency and positively correlated with demand shocks. Moreover, these studies find that $TFPR_{is}^{cs}$ dispersion is slightly lower than dispersion in technical efficiency, indicating prices are inversely related to technical efficiency under downward sloping demand.

[40] For example, on the entry/exit margins see Hopenhayn (1992) and Hopenhayn and Rogerson (1993). For adjustment cost models at the firm-level on employment, see Cooper, Haltiwanger and Wiliis (2007) and Elsby and Michaels (2013).

motivating framework. We consider the relationship between productivity, growth, and exit for all plants, exiters, and incumbents separately and estimate the following specification:

$$y_{it+1} = b_1\omega_{it} + b_2\theta_{size_{it}} + \mathbf{x}'_{it}\boldsymbol{\delta} + \epsilon_{it+1}, \tag{17}$$

where y_{it+1} denotes an outcome of plant i (growth between periods t and $t+1$ or exit), ω is plant-level productivity in logs, $\theta_{size_{it}}$ is the control for initial size (log employment) in the period, and $\mathbf{x}_{it}$ is a vector of additional controls: plant age, state effects and full interactions of industry and year effects, and the change in the state-level unemployment rate that controls for cyclical effects.[41]

Panel A of Table 3 shows $\widehat{b}_1$ from equation (17) for each outcome (column 1) and productivity estimator (columns 2-4). All point estimates are statistically significant and results from the alternative measures are similar. A one standard deviation increase in $TFPR^{cs}_{is}$ yields a roughly 4.3-log-point increase in growth and a 1.4-log-point decline in the probability of exit. For $TFPR^{rr}_{is}$, the analogous estimates are 4.7 and 1.4 log points. The implication is that productivity and growth (exit) are positively (negatively) associated, irrespective of how productivity is measured. The similarity of conclusions is one of the reasons why Syverson (2011) states that the finding that high productivity plants are less likely to exit is one of the most ubiquitous findings in the literature.[42]

Table 3: Productivity and size impact on outcomes by productivity estimator

	CS	OPH	OPHD
Panel A. Effect of productivity ($\widehat{b}_1$)			
growth	0.140***	0.136***	0.097***
exit	-0.046***	-0.042***	-0.031***
conditional growth	0.051***	0.057***	0.039***
Panel B. Effect of log size ($\widehat{b}_2$)			
growth	0.022***	0.014***	0.020***
exit	-0.034***	-0.031***	-0.033***
conditional growth	-0.051***	-0.053***	-0.051***

The table shows $\widehat{b}_1$ and $\widehat{b}_2$ in equation (17). Outcomes are employment growth among all establishments (row 1), exit (row 2), employment growth among continuers (row 3).*** denotes statistical significance at 1%. Standard errors are clustered at the state level. All regressions are based on trimmed productivity distributions (top and bottom 1% in each industry and year). Sample size information can be found in table A.1 of Appendix A.6. Source: authors' calculations.

Panel B of Table 3 illustrates the impact of initial size on outcomes. Larger business are more likely to survive and, conditional on survival, have lower growth rates. The magnitude of the size effects are similar across specifications using alternative productivity measures. Overall, we find the marginal impact of productivity on growth and survival is quite robust across different productivity

[41] This is a simplified version of the specification considered by FGH. We follow them using integrated ASM-LBD data for this analysis. The ASM provides the distribution of plant-level productivity in any given year and the LBD provides the growth and survival outcomes for the full set of plants in the ASM in that year between t and $t+1$.

[42] As in section IV.A, it is useful to highlight the role of pooling: the similarity of results conceals potentially important heterogeneity. That is, the coefficients could differ more across methods if we allowed for industry-specific heterogeneity in the effect of productivity or size.

measures even when accounting for size and age.

The findings in Sections IV.A and IV.B help explain why $TFPR_{is}^{cs}$ remains a commonly used measure of firm performance in the empirical firm dynamics literature. Since Baily, Hulten and Campbell (1992) and Foster, Haltiwanger and Krizan (2001), $TFPR_{is}^{cs}$ has been commonly used for investigating a range of issues from the determinants of firm-level growth and survival, adjustment costs for capital and labor, and the relationship between firm performance, exporter status, and management practices (see, e.g., Syverson (2011) for a survey). Ten years after HK raised questions about the interpretation of this measure, $TFPR_{is}^{cs}$ still remains commonly used in recent papers about firm dynamics.[43] One implication of our analysis is that, under our core assumptions, $TFPR_{is}^{rr}$ should be used instead of $TFPR_{is}^{cs}$ in the literature. The continuing popularity of $TFPR_{is}^{cs}$ may be partly due to the fact that, despite its theoretical appeal, $TFPR_{is}^{rr}$ estimation using a control function has practical limitations (e.g., it often requires defining industries at a 2 or 3-digit level to obtain sensible estimates for all industries) while calculating $TFPR_{is}^{cs}$ at a detailed industry level is straightforward. In light of the results in this section, we believe the findings in the literature using $TFPR_{is}^{cs}$ would likely be robust to the conceptually more appropriate index.[44] It remains an open question why these distinct measures are so tightly linked empirically (although there are competing explanations such as adjustment costs and variable markups that increase with fundamentals). It might be tempting at this juncture to argue the choice of productivity estimator is not important. However, we now turn to analysis of AE where this choice is of critical importance.

IV.C Allocative efficiency: sectoral level results

We now examine the sensitivity of allocative efficiency (AE) to the changes in the joint distribution of output and demand elasticities. Our generalization of AE is an ideal metric to assess these issues because the model requires us to take a stand on the shape of both the demand and production functions in each industry. We begin by considering simple averages of AE across industries, comparing across estimation methods. Then, in Section IV.D, we explore two further complicating factors: aggregation across industries and time-variation in parameters.

At the industry level, we need estimates of both ρ_s and γ_s to quantify sectoral $TFPQ$ and AE. To explore these issues in the context of our CS and OPH estimates, we consider the following values for the demand elasticity: $\rho_s = (0.75, 0.8, 0.94, \rho_s^{\mathrm{DW}}, \widehat{\rho}_s)$ (where $\widehat{\rho}_s$ is from the OPHD estimation). Imposing the first value across industries, paired with CS, serves as a useful benchmark because it

[43] For example, two recent prominent papers that use this as a measure of firm performance are Bloom et al. (2019) and Ilut, Kehrig and Schneider (2018). A main finding of the former paper is that plants with more structured management practices have higher $TFPR_{is}^{cs}$. The latter uses $TFPR_{is}^{cs}$ to identify nonlinear responsiveness between hires and separations at the plant-level fundamentals. Some recent papers have highlighted the distinction between measures, including Hottman, Redding and Weinstein (2016), Garcia-Marin and Voigtländer (2019), Eslava and Haltiwanger (2018), and Decker et al. (2017).

[44] One reason $TFPR_{is}^{cs}$ might be preferred is unmeasured quality differences in material inputs across plants. Inclusion of variation in output prices can help control for unmeasured input price variation since they tend to be positively correlated, see De Loecker et al. (2016).

corresponds to the approach in HK and more recently BKR.[45] Setting ρ_s=0.8 equals the mean of the ρ_s distribution obtained using the DW method. The ρ_s=0.94 case is relevant because it is equal to the mean of the OPHD-based ρ_s distribution.[46] ρ_s^{DW} and $\widehat{\rho}_s$ denote non-degenerate demand elasticity distributions implied by DW and OPHD, as in section III. The objective is to explore the sensitivity of AE to the range of demand elasticity estimates.

CS and OPH are conceptually different in the way the revenue curvature is estimated, so we describe the implications of these conceptual differences here, and organize results according to these distinct methodologies. Under CS, γ_s=1 is posited and therefore revenue curvature is fully determined by changing the value of ρ_s. Thus, the implications of different ρ_s values under CS provide guidance about the effect of changes in curvature on AE when returns to scale are fixed. Note that since cost shares are invariant to the elasticity of demand, the distribution of $TFPR_{is}^{cs}$ does not change under CS. However, the measured distribution of $TFPQ_{is}$ will change with ρ_s (e.g., the variance and correlation of $TFPQ_{is}$ with $TFPR_{is}^{cs}$).

Under OPH, revenue curvature is based on the sum of estimated revenue elasticities as described in Section III. Variation in the demand elasticity holding revenue elasticities constant yields distinct insights. First, for a given set of revenue elasticities, an increase in ρ_s implies an equivalent decline in γ_s. This enables us to evaluate the impact of decomposing the revenue curvature into its components. In addition, the distribution of distortions ($TFPR_{is}^{cs}$) may vary as ρ_s changes because the cost shares implied by β_{js} estimates may be different. Third, the distribution of $TFPQ_{is}$ will vary with the ρ_s vs. γ_s combinations for a given overall revenue curvature.

We recognize we are combining estimates from distinct estimation methodologies in this section. The motivation is to conduct a sensitivity analysis of AE to different combinations of markups and returns to scale from the range that emerge from different methods. In much of the literature calibrating AE, estimates of the requisite output and demand elasticities are from disparate sources. Still in interpreting the results it is useful to keep in mind the specifications from internally consistent approaches. That is, under the CS approach the DW estimates of ρ_s are consistent with the assumed CRS. Under the OPH approach, the OPHD estimates of ρ_s are internally consistent with the control function approach estimation.

Figure 3 shows the cross-industry average of AE_s^{COV} calculated using the range of demand elasticity estimates listed above. In order to abstract from high-frequency variation in the empirical analysis, we calculate decade-specific time series averages of AE. Several observations emerge from Figure 3. First, consider the results for CS where CRS is imposed. We find that an increase in average ρ_s yields a decline in average AE and a greater decline in average AE over time. For example, average AE declines by 20 percent from the 1970s to the 2000s under CS if ρ_s is 0.75 (for all sectors) but by 33 percent if ρ_s=0.94. Dispersion in ρ_s across industries has only a modest impact on average AE for the CS case. In many respects, the CS findings are not surprising given recent literature. For

[45]Since $\rho_s = \frac{\sigma_s - 1}{\sigma_s}$, choosing $\sigma_s = 4$ implies $\rho_s = 0.75$.

[46]Setting ρ_s=0.8 corresponds to σ_s=$(1-\rho_s)^{-1}$=5, i.e. a 25% increase in σ, while ρ_s=0.94 corresponds to an approximately 4-fold increase in σ_s.

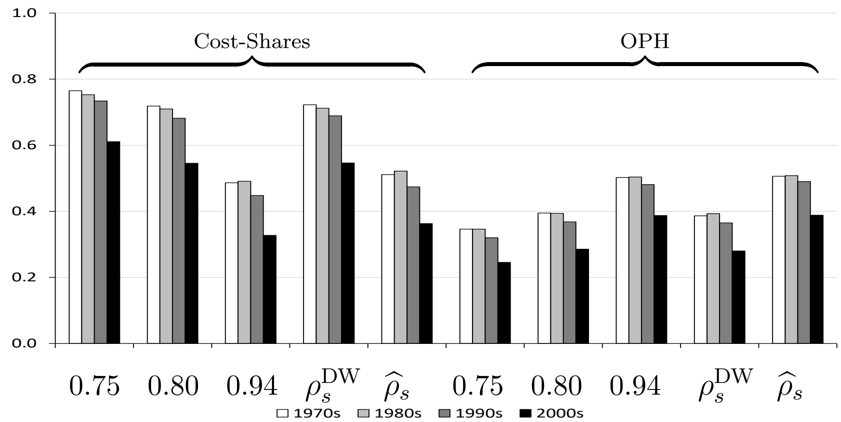


Figure 3: Unweighted average, $N_S^{-1} \sum_s \widehat{AE}_s^{COV}$

Note: The x-axis depicts alternative values of ρ_s. In industries where $1 < \rho_s\gamma_s$ at the 4-digit level, 2-digit estimates are used. ρ_s^{DW} denotes industry-specific time series averages calculated as in De Loecker and Warzynski (2012).

example, BKR indicate that under CRS, less curvature yields lower AE.

We next consider the results with OPH. Average AE is lower for OPH compared to CS for average ρ_s values substantially below one. For example, holding demand elasticities constant at the DW estimates, the CS method yields average sectoral AE that is about twice that for OPH. Recall that under OPH overall revenue curvature is determined by the estimated revenue elasticities so lower average ρ_s implies higher average γ_s for given estimated revenue elasticities. Comparing the CS and OPH results for $\rho_s = 0.75$ (for all sectors) is thus comparing results with CRS vs. generally increasing returns to scale, since overall revenue curvature is substantially below one for CS and just below one on average for OPH. As average ρ_s increases for OPH, average AE increases rather than decreases as with CS. This finding illustrates the sharp differences between CS and OPH. We find with OPH that even with overall revenue curvature constant, average AE is higher for higher average ρ_s and lower average γ_s. Thus, variation in AE is not driven simply by overall (average) curvature but by the combination of (average) ρ_s and γ_s. The combination of ρ_s and γ_s, holding the revenue curvature constant, also influences the time series decline in AE. For $\rho_s = 0.75$ under OPH, the decline from the 1970s to 2000s is 29 percent, while for $\rho_s = 0.94$ the decline is 23 percent. Allowing for heterogeneity in ρ_s for OPH does not have significant implications relative to when it is held at the cross-industry average.

Focusing on internally consistent approaches, AE using the $\widehat{\rho}_s$ under OPH is 35 percent lower on average than the benchmark CS case with a $\rho_s = 0.75$ used by HK and BKR and about 30 percent lower on average than the CS case with DW estimates of ρ_s. Thus, for the same revenue and input data, the choice of estimation method for productivity yields distinct differences in measured AE.

To help shed light on what underlies the patterns above, we return to equation (12). The second term of this decomposition shows that AE_s depends on the covariance between $TFPR_{is}^{cs}$ and

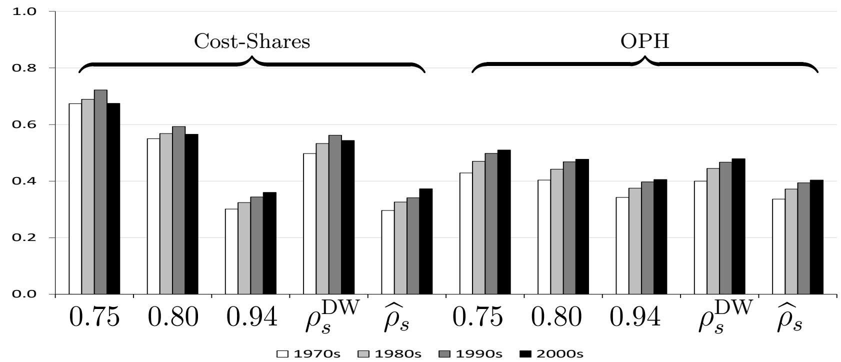
$TFPQ_{is}$. Other things equal, the second term of this decomposition implies that AE_s decreases with more dispersion in distortions ($TFPR_{is}^{cs}$), a larger positive correlation between distortions and fundamentals ($TFPQ_{is}$), and (given such a positive correlation) more dispersion in fundamentals. Figures 4(a)-4(c) show that while dispersion in log$TFPR_{is}^{cs}$ is very similar across estimation methods, dispersion in log$TFPQ_{is}$ and the partial correlation of log$TFPQ_{is}$ with log$TFPR_{is}^{cs}$ are quite sensitive to the alternative elasticity estimates. These patterns are consistent with the discussion in section I.B regarding the asymmetric impact of ρ_s and γ_s on the distribution of measured $TFPQ_{is}$.[47] Focusing on internally consistent estimation approaches, dispersion in logTFPQ$_{is}$ is notably lower and the correlation between log$TFPQ_{is}$ with log$TFPR_{is}^{cs}$ notably higher with $\widehat{\rho}_s$ under OPH compared to using either benchmark (0.75) or DW estimates of ρ_s under CS.

The patterns in Figures 4(a)-4(c) reflect the impact of the estimated parameters on the measured dispersion of fundamentals and the correlation between fundamentals and distortions. There is also a direct effect of variation in ρ_s and γ_s on AE as shown in the decomposition of AE in equation (12). Appendix A.3 illustrates the combined impact of the measurement effects and the direct effects of variation in these parameters on the first and second terms of equation (12). We find that the variation in the second term of (12) dominates the variation in AE_s empirically. Moreover, consistent with the above discussion, we find that the combined measurement and direct effects of variation in these parameters are associated with large differences in estimated AE.

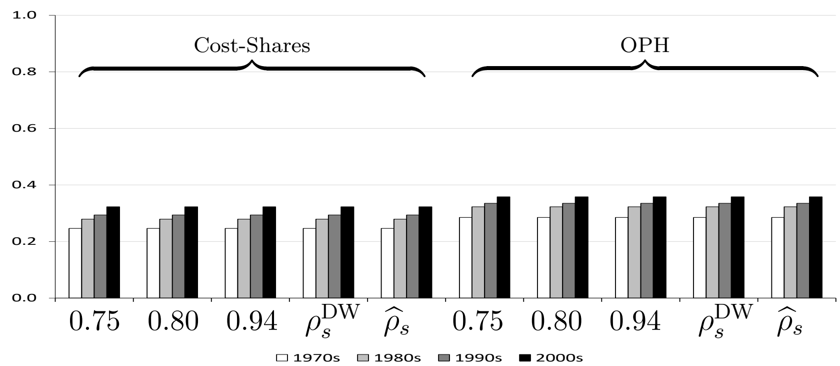
Putting these results in perspective with the findings in earlier sections, the sensitivity of measured $TFPQ_{is}$ to the estimation methods and elasticities is in contrast to the findings above that $TFPR_{is}^{rr}$ has a robust relationship with respect to $TFPR_{is}^{cs}$. The reason is that measured $TFPR_{is}^{rr}$ only depends on the overall revenue elasticities while measured $TFPQ_{is}$ depends on the decomposition of revenue elasticities into output elasticity (and hence γ_s) and demand elasticity (and hence ρ_s) components.

While this section has emphasized the sensitivity of AE to estimation methods and parameter estimates, there are some common messages from this analysis that are robust to the variation in elasticity estimates. In all specifications, measured AE is declining over time, dispersion in logTFPQ$_{is}$ and log$TFPR_{is}^{cs}$ are rising over time, and the correlation between these two alternative measures is rising over time. Rising dispersion in distortions as measured by log$TFPR_{is}^{cs}$ is one of the most robust findings across the estimation methods and parameter estimates depicted in Figure 4(b). Other things equal, this yields a decline in AE. In addition, rising dispersion of fundamentals and a rising correlation of fundamentals with distortions are working in the same direction to induce a decline in AE. Thus, a complete explanation of rising AE must account not only for the rise in the dispersion of distortions, but also a concordant increase in the correlation between distortions and fundamentals.

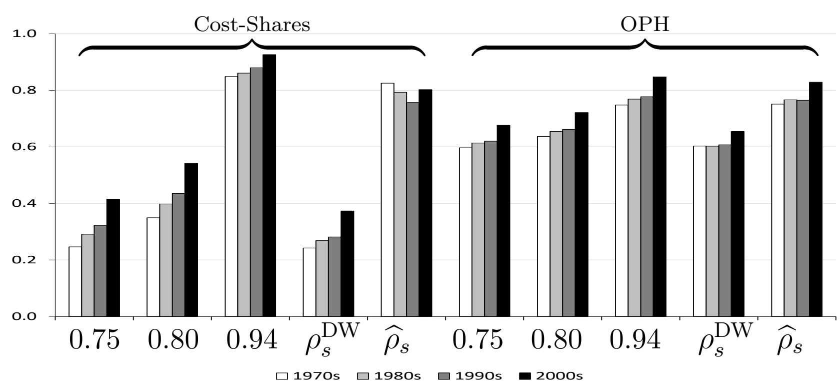
[47] By construction, there is no variation in $TFPR_{is}^{cs}$ under CS with variation in ρ_s. This also holds for OPH but log$TFPR_{is}^{cs}$ is computed differently under OPH using internally consistent cost shares cs$_{js}$=$\beta_{js}/\sum_{ks}\beta_{ks}$.



(a) Log$TFPQ$ dispersion



(b) Log$TFPR$ dispersion



(c) Partial correlation between log$TFPQ$ and log$TFPR$

Figure 4: Productivity moments

Note: The x-axis depicts alternative values of ρ_s. In industries where $1 < \rho_s\gamma_s$ at the 4-digit level, 2-digit estimates are used. ρ_s^{DW} denotes industry-specific time series averages calculated as in De Loecker and Warzynski (2012).

IV.D Allocative efficiency: Implications of time-varying elasticities and aggregation

As a further exploration of the sensitivity of AE to the estimates of key elasticities, we consider two extensions: aggregation across industries as implied by the model and variation in elasticities over time. We consider these two extensions together since they interact in interesting ways empirically. Exploring aggregation is more suitable with the full sample of manufacturing industries, especially for the roundabout production aggregation. We begin with the CS approach since this way of computing detailed industry output elasticities and the internally consistent DW approach for computing demand elasticities are straightforward for the full sample of industries. There is, however, the important limitation to the CS approach of imposing CRS. For this reason, we close this section with a discussion of the analysis of aggregation and time varying elasticities for our 50-industry sample where we can more readily relax the CRS assumption. We do so while recognizing the OPH (including OPHD) approach is less well suited to estimation at the detailed industry level for all manufacturing industries, especially if considering time varying elasticities.

For the CS analysis in this section, cost shares are calculated as time-series averages within a 4-digit SIC industry between 1972 and 1996 and another set of time-series averages within a 6-digit NAICS industry between 1997 and 2010. We find that the change in the distribution of the cost shares between the SIC and NAICS periods is small so we do not focus on that variation in this section.[48] In contrast, there has been much attention in the recent literature (e.g., De Loecker, Eeckhout and Unger (2019)) to evidence of rising markups. Figure 5 shows the implied changes in estimates of average ρ_{st} by decade under the DW methodology for all industries in the manufacturing sector. Both the unweighted and revenue-weighted means are depicted. We find evidence of declining average ρ_{st} over time. This decline yields, on a revenue-weighted (across industry) basis, an increasing average markup from about 20% in the 1980s to 33% in the 2000s. These patterns are broadly consistent with those from De Loecker, Eeckhout and Unger (2019).[49]

Figures 6(a), 6(b), and 6(c) show average sectoral AE, fixed-supply-based AE and roundabout-production-based AE under CS across different ρ_s estimates.[50] Comparing Figures 6(a) and 3 we find that the sectoral average AE results are broadly similar between the full and 50-industry samples: an increase in average ρ_s yields lower AE in both. However, average AE declines less in the full sample. For example under ρ_s=0.8, AE declines from the 1970s to 2000s by about 27% in the 50-industry

[48] See Appendix A.4 for details. It is also worth noting that the distribution of cost shares for the full sample and the 50-industry sample are quite similar.

[49] De Loecker, Eeckhout and Unger (2019) find larger average markups that increase from about 60% to 80% using Economic Census data. However, they combine both materials and labor as variable factors of production in their markup calculations. In sensitivity analysis, they find that markups computed from labor shares are higher than those from materials shares. An additional point made by Hall (2018) and Edmond, Midrigan and Xu (2019), and acknowledged by De Loecker, Eeckhout and Unger (2019), is that cost-weighted average markups increase by substantially less than revenue-weighted average markups. At the industry-level, our markups are equivalent to cost weighted means of plant-level markups.

[50] Equation (16) simplifies in the CS case under CRS. Appendix A.5 explores roundabout-production with NCRS for the 50 industry-sample.

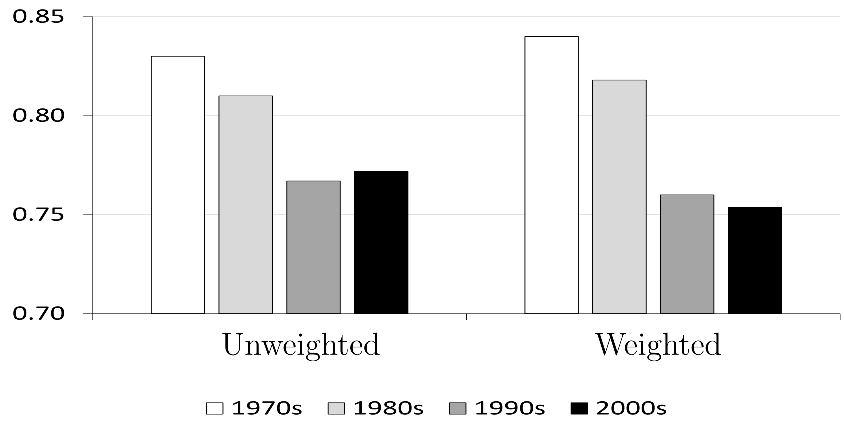


Figure 5: Descriptive statistics of ρ_{st}^{DW} estimates by decade. CS methodology, all industries.

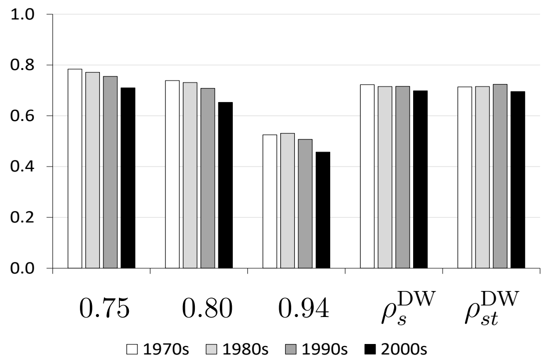
Note: "Unweighted" denotes averages where industries have equal weight. "Weighted" denotes averages where industries are weighted by revenue.

sample and about 13% in the full industry sample. In addition, sectoral heterogeneity (labeled ρ_s^{DW} in Figures 6(a) and 3) is more important in the full industry sample in that it mitigates the decline in average AE. Time series variation along with sectoral heterogeneity in ρ_s yields additional mitigation. The decline in average AE from the 1980s to 2000s with $\rho_s = 0.8$ with sectoral heterogeneity is only about 3% with sectoral heterogeneity and only about 1% with time series variation along with sectoral heterogeneity.

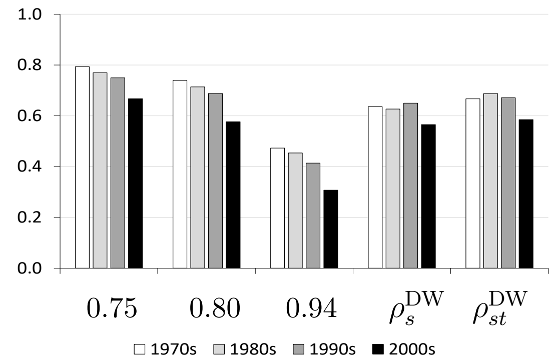
Figure 6(b) shows fixed-supply-based aggregation yields AE levels broadly similar to the sectoral average, with somewhat lower averages and steeper declines. These effects are amplified under roundabout production, shown in Figure 6(c): AE-levels are even lower and trends are even more negative. This reflects the multiplier effects of roundabout production. Heterogeneity and time variation matter here as well. Aggregate AE under roundabout production declines by 39% under heterogenous and time varying ρ_{st}^{DW} while it declines by 58% with $\rho_s = 0.8$ (which is the about the same as the mean of ρ_{st}^{DW}).

A limitation of the analysis thus far in this section is that it is restricted to the CS case under CRS. For analysis of the sensitivity to CRS, we return to our 50-industry sample and focus on non time varying elasticities. The OPHD method is not well suited to estimating time varying demand elasticities since this method exploits within industry variation over time for identification. We also focus on the revenue-weighted aggregation since most intermediate inputs are likely produced independently with respect to the 50-industry sample.[51]

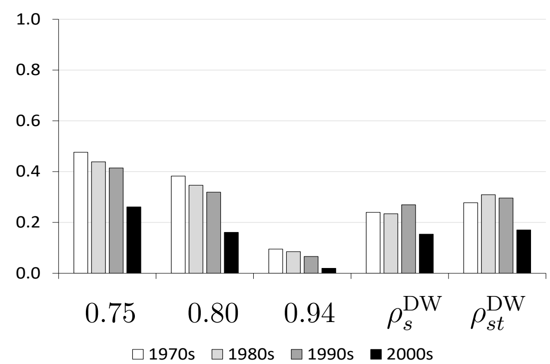
[51] Appendix A.5 presents further sensitivity analysis considering time varying elasticities and aggregation for the 50-industry sample including roundabout production using both CS and OPH. While there are some limitations of this analysis since we combine DW estimates of ρ_{st} with OPH estimates of revenue elasticities, the results show that the rise in markup over time using the DW method mitigates the decline in AE for CS as we find for the full industry sample. However, the decline in revenue elasticities using the OPH method is not as large as one would anticipate



(a) Unweighted average, $N_S^{-1}\sum_s \widehat{AE}_s^{COV}$



(b) Fixed input supply: $AE_s = \prod_s \left(\widehat{AE}_s^{COV}\right)^{\theta_s}$



(c) Roundabout production: $\prod_s \left(\widehat{AE}_s^{COV}\right)^{\frac{\theta_s}{\sum_s (1-\alpha_{Ms})\theta_s}}$

Figure 6: Descriptive statistics of AE_s^{COV} under alternative values of ρ_s (x-axis), CS methodology in full industry set

Figure 7 shows revenue-weighted aggregate AE for the 50-industry sample (for the sake of brevity we report a slightly smaller number of cases here). The patterns here are broadly similar to those for the unweighted industry means in Figure 3. However, especially for OPH, the trend declines are larger when computing the revenue weighted geometric mean across industries. Focusing on internally consistent approaches, the decline in AE using $\widehat{\rho}_s$ under OPH is more than three times larger than the decline in AE in the CS case under DW estimates of ρ_s. Thus, for the same revenue and input data, the choice of estimation method for productivity yields dramatic differences in the decline in measured aggregate AE for the 50-industry sample.

given this rise in markups. These findings suggest that there may have been offsetting increases in the returns to scale. De Loecker, Eeckhout and Unger (2019) also present some limited evidence of increases in the returns to scale over time.

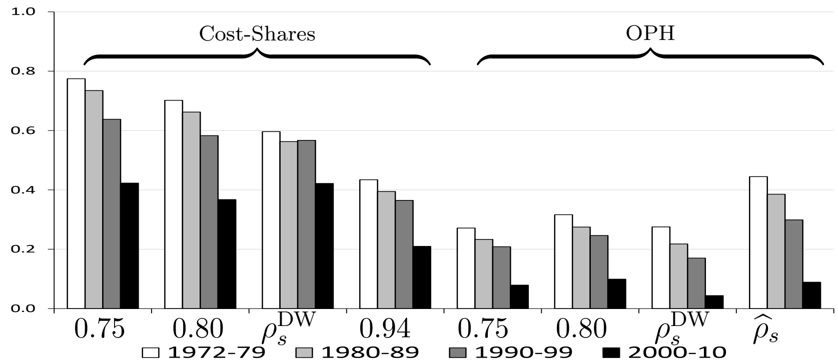


Figure 7: Fixed input supply: $AE_s = \prod_s \left(\widehat{AE}_s^{COV}\right)^{\theta_s}$

Note: The x-axis depicts alternative values of ρ_s. In industries where $1 < \rho_s\gamma_s$ at the 4-digit level, 2-digit estimates are used. ρ_s^{DW} denotes industry-specific time series averages calculated as in De Loecker and Warzynski (2012).

V Concluding remarks

When is the devil in the details of micro productivity measurement relevant for macroeconomics? We answer this question along several dimensions, investigating which details could prove devilish for researchers using micro-productivity data to inform their understanding of the economy. First, we clarify the relationship between various well-known measures of revenue productivity using HK as a framing device. Second, we draw out the implications of alternative measures for standard productivity dispersion statistics and the relationship between firm-level productivity, growth, and survival. Third, we show the importance of these measurement and estimation issues in the context of measuring allocative efficiency. An important aspect of this latter analysis is relaxing the assumption of CRS while also allowing for revenue curvature stemming from downward sloping demand. In doing so, we show that it is not just the curvature of the revenue function that matters for allocative efficiency, but also its decomposition.

Alternative productivity estimators have an important commonality: absent data on prices and quantities, they yield what have become known as revenue productivity measures. It is perhaps less recognized that the differences across estimation methods have important consequences for interpretation since the alternative measures are different conceptually. The shares of input expenditures in total costs are equivalent to output elasticities assuming CRS, while regression-based estimates are revenue elasticities absent data on prices and quantities. The revenue residuals implied by cost shares, or $\log TFPR_{is}^{cs}$, have increasingly become used as a measure of distortions. In contrast, the residual from revenue function estimation, or $\log TFPR_{is}^{rr}$, reflects fundamentals such as technical efficiency and demand shocks.

In spite of the conceptual differences between $\log TFPR_{is}^{cs}$ and $\log TFPR_{is}^{rr}$, we find that they are positively correlated, exhibit similar dispersion, and have similar relationships with firm-level growth and survival suggesting that the effect of the differences in elasticity estimation are not significant. This helps explain why log$TFPR_{is}^{cs}$ remains a commonly used measure of firm performance in the applied literature even though log$TFPR_{is}^{rr}$ is arguably conceptually the preferred measure. It remains an open question as to why these distinct measures are so tightly linked empirically although there are a number of competing explanations with empirical support (e.g., adjustment costs, variable markups that are increasing in fundamentals, or distortions that are correlated with fundamentals).

In contrast, the differences underlying these measures and estimation methods are critical for measuring allocative efficiency (AE) in a benchmark structural approach that has been developed in the recent literature. This benchmark AE is an ideal metric to assess the empirical importance of these issues because both production and demand parameters affect it directly. In addition, the joint distribution of fundamentals and distortions implied by these parameters is also critical. Our findings indicate considerable variation in both the overall revenue curvature and its components across methods. In turn, we find this yields considerable variation in the level and changes over time in measured AE. Underlying this variation is the sensitivity of the dispersion and correlation of composite measure of fundamentals, log$TFPQ_{is}$, with distortions to alternative methods. It is not only the overall curvature that matters for quantifying $TFPQ$ and AE, but also the decomposition of the curvature into its returns to scale and markup components. There is less widespread agreement across methods in this decomposition. In turn, there is less widespread agreement about the dispersion of fundamentals and the correlation of fundamentals with distortions. It is interesting that there is more widespread agreement across methods on the dispersion of distortions than there is about fundamentals themselves.

Despite the sensitivity of inferences regarding AE to estimation methodology, a common message of our findings is that all methods depict rising dispersion in both fundamentals and distortions, and rising correlation between them. All of these factors contribute to declining measured AE, which we also show may be mitigated by rising markups over time. To the extent that rising markups translate into declining revenue curvature, this mitigates the decline in AE.

One remaining (and potentially devilish) detail is the impact of heterogeneous production and demand elasticities across producers. Estimation of such elasticities is a challenge but the productivity measures we focus on are residuals that are inherently sensitive to such heterogeneity. We have taken an approach here to mitigate these concerns by using estimates of output and demand elasticities at a detailed industry level, in contrast to much of the literature that estimates key elasticities at a broader industry level. Still, detailed industry may not be sufficient. Variable markups and differences in technologies across plants are likely important contributors to measured differences in productivity across plants. This paper is intended to provide structure for thinking about these issues; it is our hope that it opens the door to future discussions.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of

Recent Production Function Estimators." *Econometrica*, 83: 2411–2451.

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. "Structural identification of production functions." University Library of Munich, Germany MPRA Paper 38349.

Baily, Martin N., Charles R. Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." In *Brookings Papers on Economic Activity: Microeconomics*. Vol. 4, , ed. Martin N. Baily and C. Winston. DC: Brookings Institute.

Baqaee, David, and Emmaneul Farhi. 2019. "Productivity and Misallocation in General Equilibrium." *Quarterly Journal of Economics*. Forthcoming.

Bartelsman, Eric J., and Wayne Gray. 1996. "The NBER Manufacturing Productivity Database." National Bureau of Economic Research Working Paper 205.

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

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2020. "Misallocation or Mismeasurement?" University of Rochester Unpublished Manuscript.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What Drives Differences in Management Practices?" *American Economic Review*, 109(5): 1648–83.

Bureau of Economic Analysis. 2015a. "BEA Fixed Asset Tables, 1972-2010." https://apps.bea.gov/national/FixedAssets/Release/XLS/Section3All_xls.xlsx.

Bureau of Economic Analysis. 2015b. "Integrated Industry Productivity Account, 1998-2010." https://apps.bea.gov//industry/iTables%20Static%20Files/AllTables.zip.

Bureau of Economic Analysis. 2015c. "Historical Integrated Productivity Industry Account, 1972-1997." https://apps.bea.gov//industry/iTables%20Static%20Files/AllTablesHist.zip.

Bureau of Labor Statistics. 2015a. "BLS Multifactor Productivity Measures, 1987-2010." https://www.bls.gov/mfp/special_requests/capmfg.xlsx.

Census Bureau. 2015a. "Longitudinal Business Database (LBD), 1972-2010." Federal Statistical Research Data Center.

Census Bureau. 2015b. "Census of Manufactures, 1972-2007." Federal Statistical Research Data Center.

Census Bureau. 2015c. "Annual Survey of Manufactures, 1973-2010." Federal Statistical Research Data Center.

Cooper, Russell W., and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies*, 73(3): 611–633.

Cooper, Russell W., John C. Haltiwanger, and Jonathan Wiliis. 2007. "Search frictions: Matching aggregate and establishment observations." *Journal of Monetary Economics*, 54: 56–78.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown." *American Economic Review*, 107(5): 322–26.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2019. "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness." Unpublished Manuscript.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica*, 79(5): 1407–1451.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review*, 102(6): 2437–2471.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2019. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics*. Forthcoming.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform." *Econometrica*, 84(2): 445–510.

Dinlersoz, Emin, and Zoltan Wolf. 2018. "Automation, Labor Share, and Productivity: Plant-Level Evidence from U.S. Manufacturing." Center for Economic Studies, U.S. Census Bureau Working Paper 18-39.

Edmond, Chris, Virgiliu Midrigan, and Daniel Xu. 2019. "How Costly Are Markups?" *Journal of Political Economy*. Forthcoming.

Elsby, Michael, and Ryan Michaels. 2013. "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows." *American Economic Journal: Macroeconomics*, 5(1): 1–48.

Eslava, Marcela, and John Haltiwanger. 2018. "The Life-cycle Growth of Plants: The Role of Productivity, Demand and Wedges." Unpublished Manuscript.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2013. "Trade and Market Selection: Evidence from Manufacturing Plants in Colombia." *Review of Economic Dynamics*, 16(1): 135–158.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics*, 34(S1): S293–S331.

Foster, Lucia, Cheryl Grim, John C. Haltiwanger, and Zoltan Wolf. 2017. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights." *National Bureau of Economic Research Working Paper Series*, No. 23666.

Foster, Lucia S., Cheryl A. Grim, John C. Haltiwanger, and Zoltan Wolf. 2016. "Firm-Level Dispersion in Productivity: Is the Devil in the Details?" *American Economic Review*, 106(5): 95–98.

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

Foster, Lucia S., John C. Haltiwanger, and Chad Syverson. 2016. "The Slow Growth of New Plants: Learning about Demand?" *Economica*, 83(1): 91–129.

Foster, Lucia S., John C. Haltiwanger, and Cornell J. Krizan. 2001. "Aggregate Productivity Growth. Lessons from Microeconomic Evidence." In *New Developments in Productivity Analysis. NBER Chapters*, 303–372. National Bureau of Economic Research, Inc.

Gandhi, Amit, Salvador Navarro, and David Rivers. 2012. "On the Identification of Production Functions: How Heterogeneous is Productivity?" Society for Economic Dynamics 2012 Meeting Papers 105.

Garcia-Marin, Alvaro, and Nico Voigtländer. 2019. "Exporting and Plant-Level Efficiency Gains: Its in the Measure." *Journal of Political Economy*, 127(4): 1777–1825.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2015. "Capital Allocation and Productivity in South Europe." *NBER Working Paper Series*, , (21453).

Hall, Robert E. 2018. "Measuring Aggregate Productivity Growth Using Plant-Level Data." *National Bureau of Economic Research Working Paper Series*, No. 24574.

Haltiwanger, John, and Zoltan Wolf. 2018. "Reallocation and Productivity: Structural Estimation and Interpretation." Unpublished Manuscript.

Haltiwanger, John, Robert Kulick, and Chad Syverson. 2018. "Misallocation Measures: The Distortion that Ate the Residual."

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): 1127–50.

Hopenhayn, Hugo A. 2014*a*. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics*, 6: 735–770.

Hopenhayn, Hugo A. 2014*b*. "On the Measure of Distortions." National Bureau of Economic Research Working Paper 20404.

Hopenhayn, Hugo A., and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy*, 101(5): 915–938.

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein. 2016. " Quantifying the Sources of Firm Heterogeneity *." *The Quarterly Journal of Economics*, 131(3): 1291–1364.

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

Hsieh, Chang-Tai, and Peter J. Klenow. 2014. "The Life Cycle of Plants in India and Mexico." *The Quarterly Journal of Economics*, 129(3): 1035–1084.

Ilut, Cosmin, Matthias Kehrig, and Martin Schneider. 2018. "Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News." *Journal of Political Economy*, 126(5): 2011–2071.

Klette, Tor Jakob, and Zvi Griliches. 1996. "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous." *Journal of Applied Econometrics*, 11(4): 343–61.

Levinsohn, James A., and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

National Bureau of Economic Research. 2014. "NBER-CES Manufacturing Industry Database, 1972-2010." http://data.nber.org/nberces/.

Olley, Steven G., and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–1297.

Peters, Michael. 2019. "Heterogeneous Markups, Growth and Misallocation." *Econometrica*. Forthcoming.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics*, 11(4): 707–720.

Syverson, Chad. 2011. "What Determines Productivity." *Journal of Economic Literature*, 49(2): 326–365.

White, T. Kirk, Jerome P. Reiter, and Amil Petrin. 2011. "Plant-Level Productivity and Imputation of Missing Data in the Census of Manufactures." Center for Economic Studies, U.S. Census Bureau Working Papers 11-02.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics Letters*, 104(3): 112–114.