Misallocation Measures:

The Distortion That Ate the Residual*

John Haltiwanger, University of Maryland and NBER

Robert Kulick, NERA

Chad Syverson, University of Chicago Booth School of Business and NBER

October 2017

Abstract. A large literature on misallocation and productivity has arisen in recent years, with Hsieh and Klenow (2009; hereafter HK) as its standard empirical framework. The framework's usefulness and theoretical founding make it a valuable starting point for analyzing misallocations. However, we show that the empirical lynchpin of this approach can be very sensitive to model misspecification. The condition in the HK model that maps from observed production behaviors to the misallocative wedges/distortions holds in a single theoretical case, with strict assumptions required on both the demand and supply sides. We demonstrate that applying the HK methodology when there is any deviation from these assumptions will mean that the "distortions" recovered from the data may not be signs of inefficiency. Rather, they may simply reflect demand shifts or movements of the firm along its marginal cost curve, quite possibly in directions related to higher profits for the business. The framework may then not just spuriously identify inefficiencies; it might be more likely to do so precisely for businesses better in some fundamental way than their competitors. Empirical tests in our data, which allow us to separate price and quantity and as such directly test the model's assumptions, suggest the framework's necessary conditions do not hold. We empirically investigate two of the possible sources of departures from the HK assumptions and implications and find support for both. We also find that measures of distortions that emerge from this approach are in fact strongly positively related with survival, suggesting they embody favorable profit conditions for the business. At the same time, however, once we condition on demand and supply fundamentals, the distortion measure becomes inversely related with survival. This suggests the measure may contain a distortionary component, but it is empirically swamped by other factors.

---

* Contact information: Haltiwanger: haltiwan@econ.umd.edu; Department of Economics, University of Maryland, College Park, MD 20742; Kulick: robert.kulick@nera.com; NERA, 1255 23rd St. NW #600, Washington, DC 20037; Syverson: chad.syverson@chicagobooth.edu; University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637. John Haltiwanger is also a Schedule A employee of the U.S. Census Bureau and Robert Kulick a Census pathways intern at the time of the writing of this paper. We thank Ron Jarmin, Pete Klenow and participants of the 2016 Summer Econometric Society Meetings, NHH, and NBER Summer Institute for helpful comments. This paper previously circulated as "Misallocation Measures: Glowing Like the Metal on the Edge of a Knife." Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

Research has established the existence of extensive heterogeneity among producers, even within narrowly defined markets. Enormous variations in establishment and firm sizes and productivity levels are ubiquitous in the data. Researchers and policymakers who focus on productivity growth have taken keen interest in the covariance of producers' size and productivity levels, because the extent to which the market succeeds in allocating activity across producers so that they are the "right" sizes (that is, they are as large as a social planner would want them to be given their relative productivity levels) affects market-, industry-, and economy-wide productivity.

A particular approach in this research genre attempts to measure "misallocations": the presence of wedges or distortions that cause producers to be either too large or too small relative to their socially efficient size. One of the seminal papers espousing this approach and introducing what has become the standard methodology for analysis of misallocations is Hsieh and Klenow (2009). The Hsieh-Klenow method combines considerable empirical power and flexibility with a straightforward measurement algorithm. From standard production microdata—revenues, along with labor and capital inputs—one can extract two producer-period-specific "wedges." One distorts the producer's input mix away from the optimal frictionless factor intensity (and through this distorts the producer's size as well), and another directly distorts the producer's size. These wedges in hand, the researcher can conduct a number of complementary empirical analyses like computing the increase in aggregate productivity if misallocations were eliminated (or brought down to some other level of interest), looking at the cross-sectional or intertemporal properties of the joint distribution of wedges, or correlating these estimated distortions with observables about the producers or the markets they operate in.

The usefulness and theoretical founding of the Hsieh and Klenow (2009) approach—hereafter HK—has driven a burgeoning and insightful literature into misallocation's productivity effects. However, we show that the empirical lynchpin of the HK approach rests on a knife's edge. The condition in the HK model that maps from observed production behaviors to the misallocative wedges/distortions holds in a single theoretical case, with strict assumptions required on both the demand and supply side. Regarding the former, every producer must face an isoelastic residual demand curve. On the supply side, producers must have marginal cost curves that are both flat (invariant to quantity) and are negative unit elastic with respect to total factor productivity measured with respect to output quantity (i.e., TFPQ).

We show that applying the HK methodology to data when there is any deviation from these elements will mean that the "wedges" recovered from the data may not be signs of inefficiency. They may simply reflect shifts in demand or movements of the firm along its (nonconstant) marginal cost curve. The producer may be employing the efficient input mix and be its optimal size, but the HK model would perceive this behavior as indicating inefficiencies. Researchers could infer misallocation when there is in fact none. What is more, under several conditions the spurious wedges actually reflect idiosyncratic demand or cost conditions that are *good* (related to higher profits) for the business. The HK method then might not just spuriously identify inefficiencies; it might be more likely to do so precisely for businesses that are in some fundamental way better than their competitors.

We go into detail below about why the production-to-wedge mismapping occurs, but we summarize it briefly now. The key implication of the HK model is that an efficient market has no variation in revenue-based total factor productivity (i.e., TFPR) among producers, even if they differ greatly in their TFPQ levels. Through the lens of the model, any observed TFPR dispersion is evidence of misallocation and the existence of distortions. This homogeneous-TFPR implication arises because in the HK model, a producer's price has an elasticity of -1 with respect to its TFPQ level. Because TFPR is the product of a producer's price and TFPQ, this negative unit elasticity ensures that TFPR is invariant to TFPQ differences across producers (or for that matter, differences over time for a given producer). For every 1% increase (decrease) in TFPQ, price falls (rises) by 1%. These changes cancel each other out, leaving TFPR unchanged. The HK model uses this invariance implication to back out misallocation measures from the TFPR dispersion that is (inevitably) observed in the data. The model reads TFPR differences as inefficiencies.

This crucial negative unit elasticity only occurs under the demand and supply conditions mentioned above: every producer must face isoelastic demand, and their marginal costs must be constant in quantity and negative unit elastic with respect to TFPQ. We demonstrate this in detail below. After demonstrating the specialness of the HK assumptions, we test whether these conditions hold in the data. We do so using a dataset where we—atypically for producer-level microdata—can observe businesses' quantities and prices separately. Specifically, we exploit the price and quantity data we developed in Foster, Haltiwanger and Syverson (2008, 2016). This allows us to directly test the model's key implication of price having an elasticity of -1 with respect to TFPQ. We find that, at least in our data spanning 11 different product markets, this condition does not hold in any market. Applying the HK framework to our data would therefore yield spurious measures of distortions.

Moreover, the elasticities of price with respect to TFPQ are consistently and considerably smaller in magnitude than one; price does not fully respond to TFPQ differences. More technically efficient businesses in our sample do not fully pass along their cost advantages to their customers through lower prices. As a result, TFPR and TFPQ are positively correlated in our sample. This positive correlation is what researchers have typically found in other samples when the data is available to compute both TFPR and TFPQ (e.g., Eslava et. al. (2013) find this using data covering all manufacturing sectors in Colombia)) and is also implied by the extensive literature on cost pass through.[1] This suggests the elasticity of price with respect to TFPQ may be less than unit elastic in magnitude more generally than just in our sample.

We conduct a second test of the implications of the HK assumptions by comparing the values of TFPQ as measured indirectly using the HK framework to the direct TFPQ measures that are feasible when price and quantity data are available. We find that the indirect measures (which we denote as TFPQ_HK) are only weakly related to the direct measures and have much higher variance. These puzzlingly findings are reconciled by using a modified HK framework with demand shocks. We find that this modified TFPQ_HK measure is much more related to demand shocks than TFPQ.

A general feature of the HK assumptions is that TFPR should exhibit no dispersion and in turn be uncorrelated with any measure of fundamentals in the absence of distortions. In a third test, we show this invariance property fails with respect to both demand- and supply-side fundamentals. One possible reconciliation of this finding is that distortions are highly correlated with fundamentals. But our analysis highlights an inherent identification problem. A strong correlation of TFPR with TFPQ and demand shocks is also exactly what one would expect if there are departures from the HK assumptions. We seek to overcome this identification problem by using our price and quantity data to directly examine these assumptions in an empirical model

[1] The positive correlation between TFPR and TFPQ in our sample is evident in Table 1 of Foster, Haltiwanger and Syverson (2008). Kulick (2016) uses the same sample for a study of horizontal mergers in ready-mixed concrete. While it is not his focus, he also finds that there is incomplete pass through of TFPQ changes on price. The broader literature on cost pass through is quite large but some examples include Goldberg and Verboven (2001); Campa and Goldberg (2005); Nakamura and Zerom (2010); Bonnet, Dubois, Villas Boas, and Klapper (2013); and Ganapati, Shapiro, and Walker (2016).

that nests, but does not impose, the HK framework. We find evidence for departures from both CES demand and constant marginal costs.

Using our more general demand and production functions, we decompose TFPR into demand and supply fundamentals as well as an alternative residual measure of distortions. This measure of distortions differs from TFPR because it accounts for variations in fundamentals like TFPQ and demand shocks that would otherwise enter into TFPR dispersion. We find that both fundamentals and the residual distortions contribute substantially to TFPR variance.

It remains unclear how to interpret our residual measure of distortions. If our more flexible demand and supply structure are still not enough to capture the full demand and cost heterogeneity in the data, those deviations would remain in the residual even though they do not indicate misallocation. (The dynamic effects of input adjustment costs, as in Asker, Collard-Wexler, and De Loecker (2014), would be one example of such deviations.) We find that, consistent with this concern, the residual measure of distortions is highly correlated with TFPR and measures of fundamentals, especially TFPQ. In addition, while TFPR, TFPQ, demand shocks, and this residual measure of distortions are all unconditionally positively associated with survival, once we control for TFPQ and demand shocks, both TFPR and the residual measure of distortions are strongly *inversely* related with survival. (TFPQ and demand shocks remain strongly positively related with survival.) This sign change of the conditional correlation suggests that TFPR and our residual measure of distortions do contain some information about factors that match the conceptualization of distortions, but these are empirically swamped by components of these distortion measures that reflect variation in fundamentals. It is only once we control for these fundamentals that the distortions are revealed. This suggests a general issue with misallocation measures: because they are essentially residuals, they may well indeed contain a kernel of distortions within them, but isolating this component from the effects of other (possibly efficient) sources of firm heterogeneity is empirically very difficult even with unusually detailed data.

The paper proceeds as follows. In section I, we review the details of the HK framework in terms of assumptions and implications. Our primary focus is to demonstrate theoretically the stringent assumptions required to use TFPR to identify distortions. Section II includes our tests of the HK assumptions and implications. Section III presents our estimates of more general demand and production function structures and quantifies their relevance for measuring distortions as well as interpreting the dispersion in TFPR. Concluding remarks are in section IV.

I. The Hsieh-Klenow Framework: Its Assumptions and Applications

A. A Brief Overview of the Hsieh-Klenow Framework

We first review the most critical elements of the Hsieh and Klenow (2009) framework. Readers seeking more detail are of course referred to the article.

The HK framework posits that each industry contains a continuum of monopolistically competitive firms (indexed by *i*) that differ in their TFPQ levels, A_i. Each firm combines labor and capital inputs to produce a single good. Firms in an industry face a Dixit-Stiglitz-type constant elasticity demand system, so each faces a residual demand curve with elasticity η. Firms choose a quantity (equivalently, price) to maximize the profit function:

$$\pi_i = (1 - \tau_{Yi})P_i Q_i - WL_i - (1 + \tau_{Ki})RK_i$$

subject to the firm's inverse residual demand curve, $P_i = Q_i^{-1/\sigma}$, and the production function $Q_i = A_i L_i^{\alpha} K_i^{1-\alpha}$.

The nonstandard elements here are the two wedges τ_{Yi} and τ_{Ki}. The former is a firm-specific scale distortion (effectively a tax or subsidy on the firm's output) and τ_{Ki} is a firm-specific factor price wedge/distortion. Their effects in equilibrium are discussed below.

Given the isoelastic residual demand curve, Firm *i*'s profit-maximizing price is then

$$P_i = \frac{\sigma}{\sigma - 1} MC_i$$

where MC_i is the firm's marginal cost, equal to

$$MC_i = \left(\frac{R}{\alpha}\right)^{\alpha} \left(\frac{W}{1-\alpha}\right)^{1-\alpha} \frac{(1 + \tau_{Ki})^{\alpha}}{A_i(1 - \tau_{Yi})}$$

The factor prices—assumed constant across firms—are *R* for capital and *W* for labor. Note that both wedges/distortions affect the firm's marginal cost and price, and firms with higher A_i (TFPQ) have lower marginal costs and prices.

At the optimal price and quantity, the firm's marginal products of labor and capital are proportional to the product of the factor price and functions of one or both distortions:

$$MRPL_i \propto W \frac{1}{1 - \tau_{Yi}}$$

$$MRPK_i \propto R\frac{1+\tau_{Ki}}{1-\tau_{Yi}}$$

Note that because of the assumption of common factor prices, in the absence of distortions, marginal revenue products of both factors would be equated across firms.

The critical result of the HK setup is that, under its assumptions, TFPR is proportional to a weighted geometric average of the marginal products of labor and capital, where the weights are the factors' output elasticities. As a result, the only firm-level variables that shift $TFPR_i$ are the two distortions:

$$TFPR_i \propto (MRPL_i)^{1-\alpha}(MRPK_i)^{\alpha} \propto \frac{(1+\tau_{Ki})^{\alpha}}{1-\tau_{Yi}}$$

This key result is what allows those who impose the HK framework to infer the presence and size of misallocations from observed differences in TFPR across producers.[2]

B. The Assumptions Driving HK's Result

The reason TFPR is invariant across firms in the HK model can be seen by recalling the definition of TFPR as the product of price and TFPQ, $TFPR_i \equiv P_i A_i$, and by substituting the expression above for the firm's marginal cost into the HK model's optimal pricing equation:

$$P_i = \frac{\sigma}{\sigma-1}\left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha}\right)^{1-\alpha}\frac{(1+\tau_{Ki})^{\alpha}}{A_i(1-\tau_{Yi})}$$

Notice that the elasticity of the firm's price P_i with respect to its TFPQ level A_i is -1. This means that as TFPQ levels and therefore prices vary across firms, the constancy of their product, TFPR, is preserved. Regardless of the characteristics of the distribution of A_i across firms, then, TFPR will not vary unless there are distortions τ_{Yi} and τ_{Ki}.

We can dig deeper into the TFPR invariance condition by using the chain rule to expand the elasticity of price with respect to TFPQ, accounting for the fact that the firm's price is a function of marginal cost, which itself depends on TFPQ. Multiplying and dividing the resulting expression by marginal cost yields (we suppress the firm index here and below when it is not necessary for clarity):

[2] This invariance of TFPR with respect to TFPQ was actually first noted by Katayama, Lu, and Tybout (2009), though they did not have distortions in their model, nor were they framing their result as being informative about misallocation. Their work points out that under their assumptions, TFPR does not reflect a firm's technical efficiency whatsoever, but rather only the factor prices it faces.

$$\varepsilon_{P,A} = \frac{dP\big(MC(A)\big)}{dA}\frac{A}{P} = \left(\frac{dP}{dMC}\frac{dMC}{dA}\right)\frac{A}{P} = -1$$
$$\frac{dP}{dMC}\frac{MC}{P}\frac{dMC}{dA}\frac{A}{MC} = -1$$
$$\varepsilon_{P,MC}\varepsilon_{MC,A} = -1$$

Equivalently,

$$\varepsilon_{P,MC} = \frac{1}{-\varepsilon_{MC,A}}$$

This decomposition of the key HK condition makes clear how the assumed functional forms on both sides of the market are necessary for the condition to hold. The elasticity of a firm's price with respect to marginal cost $\varepsilon_{P,MC}$ depends on the firm's residual demand curve, while the elasticity of its marginal cost to its TFPQ level $\varepsilon_{MC,A}$ depends on its marginal cost curve (and through this, its production function).

These demand- and supply-side components of the HK condition are not completely independent, however, because they hold at the profit-maximizing price. As such the marginal cost in the expression is evaluated at the firm's optimal quantity. This quantity depends on both the demand and cost curves. The elasticity of the firm's marginal cost with respect to TFPQ, $\varepsilon_{MC,A}$, depends both on the direct effect that TFPQ changes have on the marginal cost curve plus any movement along the marginal cost curve that a TFPQ change would induce due to a shift in the intersection between the marginal cost and marginal demand curves.

Further to this point, the $\varepsilon_{P,A}$ = -1 condition must hold at all quantities firms might produce to obtain the HK invariance condition. For example, while $\varepsilon_{P,MC}$ = 1 may hold at a particular quantity for a variable-elasticity demand system, only firms producing this exact quantity would conform to the assumptions of the HK model. All other industry firms would not, and the invariance of their TFPR levels to their TFPQ levels would not hold.

While any combination of demand- and cost-side elasticities that multiply to negative one will conform to the $\varepsilon_{P,A}$ = -1 condition, the most natural case would be where $\varepsilon_{P,MC}$ = 1 and $\varepsilon_{MC,A}$ = –1, because (as we show below) commonly assumed demand and production functions produce these results. The other cases where the product still happens to be -1 are even more "just-so" conditions than the unit elastic cases discussed here.

B.1. The Demand-Side Assumption

We now investigate the demand- and supply-side conditions under which the HK demand and cost assumptions hold. (Recalling they are connected through their evaluation at the marginal cost at the firm's profit-maximizing quantity.) We begin with demand systems where the elasticity of the firm's price with respect to its marginal cost, $\varepsilon_{P,MC}$, equals one.

When $\varepsilon_{P,MC} = 1$, the ratio of price to marginal cost is constant. That is, the price at any quantity must be a constant multiplicative markup of marginal cost, $P = \mu \cdot MC$. As is well known, this requires an isoelastic residual demand function, $Q = DP^{-\sigma}$, where D is a demand shifter and σ is the price elasticity of demand. Note that any $\sigma > 1$ is consistent with the HK assumption (the $\sigma > 1$ condition reflects the fact that profit maximization requires a firm to operate only on an elastic portion of its demand curve). As long as demand is isoelastic, it is the case that $\varepsilon_{P,MC} = 1$ regardless of the particular value of σ.

Isoelastic demand is not just consistent with the HK framework, it is the *only* form of demand that is compatible with it.[3] If firms face any other type of residual demand curve, $\varepsilon_{P,MC} \neq 1$ and the necessary condition does not hold.

To see this in an example, suppose demand is linear: $Q = a - bP$. A firm's profit maximizing price is then P = (a/2b) + (MC/2), where MC is the firm's marginal cost. (We assume MC is constant in quantity here to focus on HK's demand-side condition.) Therefore $\varepsilon_{P,MC} = (1/2)(MC/P)$. For any $P \geq MC$, $\varepsilon_{P,MC} \leq ½$. Thus with linear demand there are no situations under which the HK assumption hold, even approximately. Another illustrative example is the constant *absolute* markup demand function $Q = \lambda\exp(-P/M)$, where M is the markup. Here, $P = MC + M$ and $\varepsilon_{P,MC} = MC/(MC + M)$. In this case $\varepsilon_{P,MC} = 1$ only when the market is perfectly competitive and $M = 0$. If there is any markup, $\varepsilon_{P,MC} < 1$.

Both of these examples have the property that the elasticity of price with respect to marginal cost is always (weakly) less than one. As noted in the prior section, the results from the empirical literature suggest this property may apply more generally in the data. Previous work has typically found TFPQ to be positively correlated with TFPR, rather than uncorrelated as implied by HK. Working from the results above, this positive correlation implies that in the data the elasticity of price with respect to TFPQ is less than one in absolute magnitude:

$$\left|\varepsilon_{P,MC}\varepsilon_{MC,A}\right| < 1$$

[3] Save again for the coincidental case where a non-unitary $\varepsilon_{P,MC}$ is equal to the negative of the reciprocal of $\varepsilon_{MC,A}$ at all quantities.

Or, because theory implies $\varepsilon_{P,MC} \geq 0$ and $\varepsilon_{MC,A} \leq 0$ under standard demand and cost conditions,[4]

$$\varepsilon_{P,MC} < \frac{1}{|\varepsilon_{MC,A}|}$$

The intuition here is that for any given responsiveness of marginal costs to TFPQ, a sufficiently small pass through of lower costs (where costs reflect TFPQ) will ensure price stays high enough so that total revenues and TFPR rise when TFPQ does. Given the positive correlations found in empirical work, this smaller pass through appears to be the typical case in the data.

B.1. The Supply-Side Assumption

We now consider the supply-side necessary condition for HK's result: the elasticity of the firm's marginal cost at its optimal quantity with respect to its TFPQ level is negative one. By definition, this holds when

$$\varepsilon_{MC,A} = \frac{\partial MC\big(A, Q(A)\big)}{\partial A}\frac{A}{MC} = -1$$

where $MC(A,Q(A))$ is the firm's marginal cost function (the derivative of its cost function with respect to quantity). We have explicitly written the firm's quantity as a function of TFPQ, but have suppressed the other arguments of the marginal cost function such as factor prices because they are assumed constant across firms in the HK framework.

To explore the theoretical conditions under which $\varepsilon_{MC,A} = -1$ might hold, consider first how a change in TFPQ would qualitatively affect a firm's realized marginal cost. The total change in marginal cost depends both on the direct negative effect of TFPQ on costs—the shift in the marginal cost curve—as well as any change in marginal cost resulting from the effect of TFPQ on the firm's optimal quantity—movement along the marginal cost curve. As noted above, this total effect of a TFPQ increase is bounded from above by zero (the case under perfect

[4] For smooth demand curves (those with continuous marginal revenue curves), price weakly rises with marginal cost because an increase in marginal cost reduces the firm's optimal quantity, running up the marginal revenue and demand curves. The limit case is perfect competition, where the residual demand and marginal revenue curves are flat, and a change in the firm's marginal cost has no effect on price. The change in a firm's marginal cost resulting from a change in its TFPQ level *A* depends both on the direct negative effect of TFPQ on costs and any change in marginal cost resulting from the effect of TFPQ on the firm's optimal quantity. As detailed below, this total change is weakly negative, with again the limit case being perfect competition. In that boundary case, realized marginal cost remains at the (unchanged) market price and the product of the demand- and supply-side elasticities remains less than one, though of course the second inequality is undefined in this case.

competition), which requires upward-sloping marginal cost curves. The sum of these two effects—reinforcing if marginal costs decline in quantity, countervailing if they rise—must be negative unit elastic to conform to the HK model.

The simplest case where this holds is when the marginal cost curve is flat and marginal costs are negative unit elastic in TFPQ; that is, when the marginal cost curve has the form:

$$MC(A) = \frac{\Phi(\boldsymbol{W})}{A}$$

where $\Phi(\boldsymbol{W})$ is a function of the vector of factor prices $\boldsymbol{W}$. The firm's quantity is not an argument in this function, indicating constant marginal costs in quantity. Intuitively, the negative unit elasticity holds in this case because there is no reinforcing or countervailing effect of TFPQ on the firm's optimal quantity. The only influence TFPQ has on marginal cost is its direct effect, which is negative unit elastic.

We can integrate with respect to Q to find the cost functions that satisfy the condition:

$$C(A,Q) = \int_0^Q \frac{\Phi(\boldsymbol{W})}{A} dQ = \frac{Q}{A}\Phi(\boldsymbol{W}) - F$$

where F is a fixed cost. Some commonly used cost functions have this form. For example, the Cobb-Douglas production function $Q = AL^{\alpha}K^{\beta}$ has a cost function equal to

$$C(A,Q) = \left(\frac{Q}{A}\right)^{\frac{1}{\alpha+\beta}} \left(\frac{\alpha+\beta}{\alpha^{\alpha}+\beta^{\beta}}\right)^{\frac{1}{\alpha+\beta}} W^{\frac{\alpha}{\alpha+\beta}} R^{\frac{\beta}{\alpha+\beta}}$$

As is obvious from inspection, this has the required form if $\alpha + \beta = 1$; i.e., the production function exhibits constant returns to scale. This is the production function and parameterization HK assumes.[5,6]

[5] A similar result holds for the general CES production function $Q = A[\alpha L^{\rho} + \beta K^{\rho}]^{\frac{\nu}{\rho}}$, where ρ parameterizes the elasticity of substitution between inputs and ν parameterizes the scale elasticity. In this case, the corresponding cost function is:

$$C(A,Q) = \left(\frac{Q}{A}\right)^{\frac{1}{\nu}} \left[\alpha^{\frac{1}{1+\rho}} W^{\frac{\rho}{1+\rho}} + \beta^{\frac{1}{1+\rho}} R^{\frac{\rho}{1+\rho}}\right]^{\frac{1+\rho}{\rho}}$$

Again if the production function exhibits constant returns to scale (i.e., $\nu = 1$), marginal costs will be constant and negative unit elastic with respect to TFPQ.

[6] Note that the HK framework admits nonconstant returns to scale arising from fixed costs. However, as noted by Foster et al. (2017), in practice this will rely on the empiricist being able to measure the true, marginal A_i. If instead TFP is measured according to the common practice of taking a ratio of output to weighted inputs, this ratio will not be invariant to the firm's optimal quantity, and again the HK assumptions will be violated.

However, the HK requirement that $\varepsilon_{MC,A}$ = -1 will not hold without constant returns to scale. With nonconstant returns, the effect of TFPQ on marginal costs is not just the direct effect through shifting the marginal cost curve but also the induced movement along the curve because the firm's optimal quantity changes when TFPQ does. The size of this quantity change depends on the relative slopes of both the marginal cost and marginal revenue curves around the location of the quantity change, but in any case the induced shift will not generally lead to a negative unit elastic response of marginal cost.

To see this in an example, suppose that the firm faces an isoelastic residual demand curve $Q = DP^{-\sigma}$ as the HK methodology assumes. Its inverse marginal revenue curve is then $MR = D_1 Q^{-\frac{1}{\sigma}}$, where D_1 is a constant.

Now suppose the firm's cost function has the generalized form

$$C(A,Q) = \left(\frac{Q}{A}\right)^{\frac{1}{\nu}} \Phi(\boldsymbol{W})$$

where ν parameterizes the scale elasticity (ν = 1 implies constant returns to scale in the production function). Marginal costs are $MC = \frac{1}{\nu} Q^{\frac{1}{\nu}-1} A^{-\frac{1}{\nu}} \Phi(\boldsymbol{W})$.

The firm's optimal quantity equates MR and MC. Equating the logs of these two values and solving for the firm's logged optimal quantity, we have:

$$q = \left(\frac{\nu\sigma}{\sigma - \nu - \nu\sigma}\right)\left[d_1 - c_1 + \frac{1}{\nu} a\right]$$

where lowercase q and a denote logs of quantity and TFPQ and d_1 and c_1 are respectively demand- and cost-side constants. Substituting this back into the expression for logged marginal costs, we have

$$mc = c_1 + \left(\frac{1}{\nu} - 1\right)\left(\frac{\nu\sigma}{\sigma - \nu - \nu\sigma}\right)\left[d_1 - c_1 + \frac{1}{\nu} a\right] - \frac{1}{\nu} a$$

Therefore the elasticity of marginal cost with respect to TFPQ is

$$\varepsilon_{MC,A} = \left(\frac{1}{\nu} - 1\right)\left(\frac{\nu\sigma}{\sigma - \nu - \nu\sigma}\right)\frac{1}{\nu} - \frac{1}{\nu} = -\frac{1}{\nu - \sigma + \nu\sigma}$$

This has the HK-required value of -1 when $\nu - \sigma + \nu\sigma = 1$. Solving for ν:

$$\nu = \frac{1 + \sigma}{1 + \sigma} = 1$$

Thus regardless of the slope of the residual demand and marginal revenue curves $-1/\sigma$, the only scale parameter consistent with the HK condition is ν = 1. Any nonconstant marginal costs, whether increasing or decreasing, will violate the necessary condition because the reinforcing or countervailing effect of TFPQ moving a firm along its marginal cost curve will make the elasticity of the firm's realized marginal cost with respect to its TFPQ different from -1.

C. A Graphical Demonstration of the Uniqueness of the HK Assumption

In this section, we use a graphical framework to explain why the HK framework delivers the TFPR invariance result, and why any departure from either its demand- or supply-side necessary assumptions will lead TFPR to differ across firms even if there are no distortions. This will reinforce the analysis above.

However, there is an additional point to our exercise here. We introduce firm-specific demand shifts, which are not in the baseline HK model, into the framework. We show that under the assumptions of the HK model, the demand shocks do not affect the key TFPR invariance implication. If any of the component assumptions fail, however, firm-specific demand shocks will create variation in TFPR even in the absence of distortions. This creates a second channel through which applying the HK condition can yield spurious distortion measures.

We start our analysis by imposing the HK assumptions. Residual demand is isoelastic, $Q = DP^{-\sigma}$. The corresponding inverse demand is $P = D^{\frac{1}{\sigma}}Q^{-\frac{1}{\sigma}}$ and the inverse marginal revenue curve is $MR = \left(1 - \frac{1}{\sigma}\right)D^{\frac{1}{\sigma}}Q^{-\frac{1}{\sigma}}$. Both of these curves are log-linear:

$$p = \frac{1}{\sigma}d - \frac{1}{\sigma}q$$

$$mr = \ln\left(1 - \frac{1}{\sigma}\right) + \frac{1}{\sigma}d - \frac{1}{\sigma}q = \ln\left(1 - \frac{1}{\sigma}\right) + p$$

where lowercase letters are logged values. (Neither function is defined at its vertical or horizontal intercepts.)

Because $\sigma > 1$, the first term in the logged marginal revenue curve is negative. Thus in logged-quantity-logged-price space, the marginal revenue curve runs parallel to the demand curve at a distance $\ln(1 - 1/\sigma)$ below it. As we will see below, this parallelism is important to the HK result.

We also impose the HK assumption of constant returns to scale in the production function. The corresponding cost function is

$$C(A,Q)=\frac{Q}{A}\Phi(W)$$

Marginal costs of course do not depend on output, and their elasticity with respect to TFPQ is -1. The log of marginal cost is:

$$mc=\phi(w)-a$$

These elements—the demand curve, the marginal revenue curve, and the marginal cost curve—are combined in the solution to the standard monopolist's price/quantity problem in Figure 1. The firm's optimal (logged) quantity is where $mr = mc$, q^*, and its optimal price is p^*.

The figure also demonstrates how a change in (logged) TFPQ, a, affects the optimal quantity and price. The HK condition requires that TFPR, which is the product of P and A, be invariant to changes in A. In the logged space shown in the figure, it means that any change in TFPQ, Δa, must induce a price change $\Delta p = -\Delta a$.

Figure 1 makes clear why this result always holds in the HK setting. Suppose TFPQ rises from a to a', so $\Delta a = a' - a$. HK's assumed $\varepsilon_{MC,A} = -1$ implies that $\Delta mc = -\Delta a$. This drop in marginal cost raises the firm's optimal quantity to q'^* with a corresponding price change from p^* to p'^*, as shown in the figure. Here is the key result: because the marginal revenue and demand curves $mr(q)$ and $p(q)$ are parallel and the marginal cost curve horizontal, it must be that the drop in logged marginal revenue at the optimum quantity must exactly equal the drop in logged price. Thus $\Delta p^* = \Delta mr^* = \Delta mc^* = -\Delta a$, the HK result.

Note that both elements of the HK framework are necessary for this result. Only isoelastic demand creates parallel demand and marginal revenue curves. This ensures a given change in logged marginal revenue at the optimal quantity translates into the same-sized change in logged price. In other words, the ratio of (the level of) price to (the level of) marginal cost stays the same, so the elasticity of price with respect to marginal cost is one. The constant returns assumption creates the horizontal marginal cost curve. This ensures that the total effect on the firm's marginal cost at its optimal quantity, Δmc^*, is only the direct effect of the shift in the curve Δa; there is no reinforcing (if the marginal cost curve is downward sloping) or countervailing (upward sloping) effect on marginal costs through induced shifts along the marginal cost curve when the firm's optimal quantity changes.

Violating either of these conditions ensures that $\Delta p \neq -\Delta a$ and failure of TFPR invariance with respect to TFPQ.

It is obvious from inspection of Figure 1 that any other demand curve, because it does not have a parallel marginal revenue curve, will cause any change in logged marginal cost—even in the presence of a horizontal marginal cost curve—to lead to a disproportionate change in the firm's optimal price. (Recall that proportionalism in levels is graphically reflected in parallelism in logged values.)

Regarding the HK assumption about the marginal cost curve, Figure 2 preserves CES demand but shows the effect of an increase in TFPQ when marginal costs rise with output. As in Figure 1, an increase in logged TFPQ from a to a' shifts down the marginal cost curve by Δa. Here, however, because the marginal cost curve is not horizontal, the effect of this TFPQ change on the firm's marginal cost is not just the drop in the *mc* curve. It is also the effect of moving along the new *mc* curve from the old optimal quantity q^* to the new one q'^*. This total effect is necessarily less than Δa because *mc* is upward sloping. As a result, price doesn't fall as much as the marginal cost curve shifts down, and $\Delta p \neq -\Delta a$. Similarly, a downward-sloping marginal cost curve would create a movement along the *mc* curve that would make the total effect of a change in TFPQ on marginal costs *greater* than Δa. Again, it is the case that $\Delta \mathrm{p} \neq -\Delta \mathrm{a}$.

II. Testing the Assumptions of the Hsieh-Klenow Framework

A. Elasticity of Prices with Respect to TFPQ

We first test the core implication of the HK setup: producer prices are negative unit elastic with respect to TFPQ levels.

One needs to observe prices and TFPQ levels to conduct this test. While techniques have been developed to back out otherwise unobservable price and quantity information from revenue data (see, e.g., Klette and Griliches, 1996; Katayama, Lu, and Tybout, 2009; De Loecker and Warzynski, 2012), these require assumptions, making any test a joint test not only of the assumptions of the HK model but these techniques as well.

Fortunately, we collected a dataset in earlier work (Foster, Haltiwanger, and Syverson, 2008, 2016) that includes separate quantity and price information at the individual producer level. Those papers extensively detail this data, so we only very briefly review its contents here.

Our microlevel production data is a subset of the 1977, 1982, 1987, 1992, and 1997 U.S. Census of Manufactures (CM). The CM collects information on plants' shipments not just in the standard revenue sense (i.e., dollar values), but physical units as well. The sample includes producers of one of eleven products: corrugated and solid fiber boxes (which we will refer to as "boxes" from now on), white pan bread (bread), carbon black, roasted coffee beans (coffee), ready-mixed concrete (concrete), oak flooring (flooring), gasoline, block ice, processed ice, hardwood plywood (plywood), and raw cane sugar (sugar).[7] We chose these products based in part on their physical homogeneity, which allows plants' output quantities and unit prices to be more meaningfully compared.

From these product-level revenue and physical quantity data, we can construct important inputs to our analyses here. (The details of construction can be found in our earlier work.) First, we can compute producers' average unit prices. Second, we can measure TFPQ directly, using physical quantity as the output measure in the productivity numerator. Third, we can back out idiosyncratic demand shifts (alternately referred to as "shifts" and "shocks" below) for every producer. We describe this process in Foster, Haltiwanger, and Syverson (2008), but in brief, we impose a CES demand system for each industry—using TFPQ as a cost-shifting instrumental variable—and take the residual as a measure of the producer-specific demand shift.

In our basic specification, we regress a producer's logged price on its contemporaneous logged TFPQ for each product separately:

$$p_{it} = \alpha_0 + \alpha_1 tfpq_{it} + \eta_t + \varepsilon_{it}$$

where η_t is a fixed effect corresponding to the CM year, which removes any shifts in prices across time that are common across all producers. Under the HK assumptions, $\alpha_1 = -1$. We therefore test industry-by-industry the null hypothesis that $\alpha_1 = -1$.

In addition to these industry-specific tests, we estimate a pooled specification on the combined dataset. Here the specification is the same, except rather than just having CM year fixed effects we include industry-CM year fixed effects, so all identification of the relationship between price and TFPQ still comes from within-industry-year variation. Of course in this case we are imposing a common value of α_1 across all industries for the relationship between logged price and logged TFPQ.

---

[7] We exclude observations with imputed physical quantity data. For this purpose, we take advantage of newly recovered item impute flags developed and described in White, Reiter and Petrin (2014). We use inverse propensity score weights in our analysis to deal with possible non-randomness in the likelihood of observations being imputed. We find that results are largely robust to not using such weights. We use the same approach as in Foster, Haltiwanger and Syverson (2016) for this purpose. See the latter paper for details.

The results are shown in Table 1. The magnitudes of the estimated elasticities α_1 are considerably less than one for every industry. The null hypothesis of the HK conditions is clearly rejected; the smallest *t*-statistic rejecting the null is 4.4, for carbon black. In the pooled specification, we estimate an average elasticity of price with respect to TFPQ of -0.450, and reject the null with a *t*-statistic of 86.4. Thus the average elasticity of a producer's price with respect to its TFPQ level is less than half the magnitude of that implied by the HK assumptions. Given that price is less than negative unit elastic with respect to TFPQ, this means that in our data TFPR is positively correlated with TFPQ. Producers with low costs (high TFPQ) do not fully pass onto consumers their cost advantages.

We estimate two more pooled specifications as a check because, as noted in Foster, Haltiwanger, and Syverson (2008), the fact that we measure unit prices as the quotient of reported revenues and physical quantities means that measurement error in quantities might create division-bias-based measurement error in a regression of price on TFPQ. We therefore employ two instrumental variables strategies discussed in detail in our prior work. In one specification, we instrument for logged TFPQ with the producer's innovation in TFPQ from the previous CM. In the second, we instrument using the producer's TFPQ level in the previous CM. The first stage results in both cases indicate these instruments have considerable explanatory power with respect to current TFPQ. The results of the second stage, shown in Table 1, are consistent with the OLS results. In both cases, the point estimates of the elasticity $\varepsilon_{P,A}$ are well below one, economically and statistically.

In sum, we find consistent evidence that the elasticity of price with respect to TFPQ is well below one in magnitude for the producers in our data. Of course, this result applies to our particular sample, which is by no means representative of all production settings. It cannot elucidate whether the less than complete response of prices to TFPQ-driven cost changes holds more generally. On this point, however, we can point to a separate and very large empirical literature on pass through rates that indicates our result is indeed typical. Some examples of this literature include Goldberg and Verboven (2001); Campa and Goldberg (2005); Nakamura and Zerom (2010); Bonnet, Dubois, Villas Boas, and Klapper (2013); and Ganapati, Shapiro, and

Walker (2016). These studies reflect the results found in the vast majority of that literature: across diverse market settings, pass through of costs into prices is less than one-for-one.[8]

B. Relationship between Direct TFPQ Measures and TFPQ from HK Framework

We conduct a second test of the HK framework using our sample of homogenous-product manufacturers. Namely, we back out the TFPQ implied by the HK model from our data and compare it to the TFPQ that we can measure directly. This gives us the ability to gauge how closely a key unobservable backed out from the HK framework resembles its direct measure.

As HK show, one can recover a producer's implied TFPQ as follows:

$$TFPQ\_HK_i = \kappa \frac{(P_i Q_i)^{\frac{\sigma}{\sigma-1}}}{K_i^{\alpha} L_i^{1-\alpha}}$$

Intuitively, the numerator is output as backed out from observed revenue via the demand elasticity σ. We allow the elasticity to vary by industry, using our industry-specific demand estimates (described in more detail below). The denominator is the standard composite TFP input. The constant κ is the same across all producers, so it washes out in all the comparisons we make below and can be ignored.

We compare this to our directly measured TFPQ:[9]

$$TFPQ_i = \frac{Q_i}{K_i^{\alpha} L_i^{1-\alpha}}$$

It is apparent that the model-driven transformation from revenue to implied output is what separates TFPQ_HK from TFPQ.

Our data reveal that this indirect-versus-direct distinction in quantity measurement makes a big difference. The correlation between TFPQ_HK and TFPQ across our entire sample is only 0.09. That is, the physical efficiency of producers in our data as backed out from the HK model is weakly correlated with its directly measured value. Part of this poor fit reflects the fact that

[8] Note that constant-elasticity demand implies complete pass through of *logged* costs into *logged* prices. With a markup, therefore, the pass through of cost *levels* into price *levels* will be greater than one-to-one. Some of the cited studies measure pass through in levels rather than logs. Given that they find less than one-to-one pass through in levels, this also implies less than one-to-one pass through in logs. Indeed, one notable "exception" paper in the literature known for finding close to complete pass though in its empirical setting is Fabra and Reguant (2014). However, their result of near-complete pass through is in levels, indicating incomplete pass through in logs.

[9] In practice we use a gross output production function obtain TFPQ. We present a version based on a value added production function here in order to match the notation in HK (2009). Note also that we assume the production function has constant returns to scale, just as the HK framework.

there is much more variability in TFPQ_HK than TFPQ. The standard deviation of TFPQ is 0.28, while for TFPQ_HK it is an enormous 3.29.[10] The source of this large variance can be observed in the expression for TFPQ_HK above: a demand elasticity σ near one requires huge variation in implied quantity to explain observed revenue variation. Two of our sample industries, carbon black and gasoline, have estimated demand elasticities that are relatively close to unity and as such have highly variable implied output quantities. If we remove these from the sample, the standard deviation of TFPQ_HK falls considerably, to 1.03. However, this is still much larger than the TFPQ standard deviation for this restricted sample of 0.28, and in any case the main message stands: TFPQ_HK and TFPQ are only weakly correlated, with a correlation coefficient of 0.29 in this restricted sample.

At the same time, TFPQ_HK is uncorrelated with producers' prices (correlation coefficients of 0.01 in the whole sample and 0.01 in the sample excluding carbon black and gasoline). This contrasts with a correlation between directly measured TFPQ and prices of -0.59 in both the whole and restricted samples. As Foster, Haltiwanger, and Syverson (2008) point out, this negative correlation is consistent with the notion that TFPQ differences are cost differences: higher TFPQ implies lower costs, and these costs are then (partially) passed through in the form of lower prices. The fact that increases in TFPQ_HK do not correspond to lower prices raises questions about the extent to which TFPQ_HK captures firms' cost efficiencies.

We next compare TFPR to alternative measures of TFPQ and our measured estimated producer-level shifts. TFPR has slightly lower dispersion (standard deviation of 0.23) than directly measured TFPQ, but it is much less dispersed than TFPQ_HK. TFPR is highly correlated with directly measured TFPQ (about 0.66) and positively correlated with demand (0.29). On the other hand, it less correlated with TFPQ_HK (0.11). These patterns are for the full sample but are quite similar for the restricted sample.[11] This further emphasizes the point that measures of TFPQ derived from the HK model do not behave the way directly measured TFPQ does in our sample.

[10] These calculations use values where we have removed industry-year means from the sample. Bear in mind that these TFPQ values are in logged units of output, so a log difference of 3.29 implies a 27-fold ratio in levels.

[11] The correlation between TFPR and TFPQ_HK rises to 0.36 in the restricted sample. The other correlations are very similar to the full sample. Both Hsieh and Klenow (2009) and Bils, Klenow, and Ruane (2017) report a low correlation between TFPR and TFPQ_HK, about 0.10.

One potential source of the unusual relationships between TFPQ_HK and directly measured TFPQ and prices is that we apply the baseline HK model to derive TFPQ_HK (excepting the fact that, unlike HK, we use industry-specific demand elasticities). The baseline model has no demand shifts across producers; all heterogeneity comes through TFPQ and distortions. However, it is possible—and indeed a burgeoning literature suggests it is likely—that producers face idiosyncratic demand shocks along with having different productivity levels. Hsieh and Klenow (2009) show in an appendix that their model can be augmented to include demand shocks (horizontal shifters in firms' CES demand curves) while still preserving the basic logic of the model. To see how allowing demand variations might improve the fit of TFPQ_HK to TFPQ, we apply this augmented version of their model to our data.

In the demand-augmented HK framework, TFPQ_HK is now[12]

$$TFPQ\_HK\_WD_i = \kappa \frac{(P_i Q_i)^{\frac{\sigma}{\sigma-1}}}{K_i^{\alpha} L_i^{1-\alpha}} = \kappa \frac{Q_i}{K_i^{\alpha} L_i^{1-\alpha}} D_i^{\frac{1}{\sigma-1}}$$

where D_i is firm *i*'s idiosyncratic demand. We mnemonically name the object TFPQ_HK_WD to denote "with demand." Intuitively, this is a composite measure reflecting $TFPQ_i = \frac{Q_i}{K_i^{\alpha} L_i^{1-\alpha}}$ and idiosyncratic demand. Decomposing this composite into its demand and TFPQ components is not feasible with standard production data with revenue and inputs. We are able to conduct this decomposition into TFPQ and D_i in our data, however. Under the HK assumption of a CES demand system, the estimated D_i by design satisfies the above composite shock relationship. A critical point to emphasize is that the equivalence between TFPQ_HK and the composite shock requires estimating the demand elasticities and demand shocks in an internally consistent manner. In practice, HK and others who implement the TFPQ_HK methodology typically impose the same elasticities across industries (and countries).

We find that TFPQ_HK has a stronger correlation with our measure of demand (about 0.28 in the full sample and 0.55 in the restricted sample) than with TFPQ. This suggests that a considerable amount of the variation in TFPQ_HK is actually driven by demand shifts rather than TFPQ differences.

[12] The D_i shifter we are now including is from the specification $Q_i = D_i P_i^{-\sigma}$, so that $P_i = D_i^{\frac{1}{\sigma}} Q_i^{-\frac{1}{\sigma}}$. That is, it is the shift in quantity demanded Q_i holding price constant. In this specification $(P_i Q_i)^{\frac{\sigma}{\sigma-1}} = D_i^{\frac{1}{\sigma-1}} Q_i$

To add further insights, our final investigation of the properties of TFPQ_HK explores its relationship with survival and compares this to other producer-level metrics. Table 2 shows the results. High TFPR, high directly measured TFPQ, and high demand D_i are each negatively associated with exit. Demand shocks play the dominant quantitative role, with a one standard deviation increase in demand associated with a 6-percentage-point drop in the probability of exit. In contrast, a one standard increase in TFPQ is tied to a decline in the probability of exit of 1 percentage point. High TFPQ_HK plants are also more likely to survive. A one standard deviation increase in TFPQ_HK corresponds to a 2.6-percentage-point decline in the exit probability. Overall, then, the most important predictor of exit is the demand shift, with a one standard deviation increase yielding a drop in the exit rate that is about six times larger than that of a similar sized shift in TFPQ, and more than twice that of TFPQ_HK.

In short, our evidence suggests TFPQ_HK is best thought of as a composite measure that reflects both TFPQ and demand shocks. While it is interpretable as a composite, it has less predictive value in accounting for key outcomes like survival than its underlying components. Moreover, this composite interpretation requires that demand be estimated in a manner that is internally consistent with the micro data.

C. Demand Variations and the Hsieh-Klenow Framework

The analysis in the previous section makes one thing clear: demand variations across producers are important. We explore the empirical relationship between TFPR and demand in this section, but we first discuss how demand variations fit into the HK framework more generally.

Under the joint assumptions of isoelastic demand and constant marginal costs, shifts in a firm's residual demand curve will not change its TFPR level in the absence of distortions. The inverse is also true: if either or both of these assumptions do not hold, variation in demand will create variation in TFPR.

The invariance of a firm's TFPR to demand shifts under the HK conditions is shown in Figure 3. Initially, the firm's demand and marginal revenue curves are $p(q)$ and $mr(q)$, and the firm's optimal price and quantity are p^* and q^*. The inverse demand curve then shifts by Δd. This shifts out marginal revenue by Δd as well. As a result, the firm's profit-maximizing quantity rises to q'^*. However, the profit-maximizing price remains p^*. Because the firm's price does not change and TFPQ is unaffected by the demand shift, TFPR does not change.

The intuition for this result is straightforward. Isoelastic demand implies a constant multiplicative markup. Thus if marginal cost does not change when demand shifts, price won't either. When the marginal cost curve is flat as in the HK model, shifts in a firm's demand that are uncorrelated with shifts in its marginal cost curve will not change the firm's optimal price. As a result TFPR does not move with demand. The same implication holds across firms. Differences in demand that are not correlated with TFPQ differences will not create price, and therefore TFPR, variation.

To see how departures from the HK assumptions cause TFPR to be correlated with demand even in the absence of distortions, consider the cases in Figure 4. Panel A shows an example of a non-isoelatic residual demand curve but constant marginal costs. A shift in the firm's residual demand by Δd no longer creates a parallel shift in the marginal revenue curve because the markup varies with quantity. As a result, even though marginal costs are constant, the markup, and hence price, is not. The change in price changes TFPR. Thus demand shifts TFPR if demand is not isoelatic.

In Panel B, demand is again isoelastic, but marginal costs are no longer constant. Instead the firm's marginal cost rises with its quantity. As opposed to the HK case in Figure 3, here a demand shift changes not just the firm's optimal quantity but its price too. The multiplicative markup has not changed, but the firm's marginal cost has because of nonconstant returns. As a result, the demand shift changes TFPR. Here TFPR increases with a positive shift in demand; TFPR would fall if the marginal cost curve were downward sloping.

The comparison of Figures 3 and 4 suggests a test. If one can measure demand shifts (either across firms or within firms over time) that are orthogonal to TFPQ variations, one can see if these demand changes are correlated with TFPR levels. Rejecting the null hypothesis of no correlation would indicate that either the HK assumptions do not hold or that the distortions are correlated with demand. Because the invariance of TFPR to demand changes depends on prices being invariant to demand, a corollary test that we conduct is to see if demand changes are correlated with plant-level prices.

We begin with the demand shifts we used in the prior section to explore the properties of TFPQ_HK_WD. For each product we estimate the simple specification

$$tfpr_{it} = \beta_o + \beta_1 demand_{it} + \eta_t + \varepsilon_{it}$$

where $tfpr_{it}$ is (log) TFPR for plant at time t, $demand_{it}$ is the idiosyncratic demand shift identified as described above, η_t is a CM year fixed effect, and ε_{it} is the residual. We also estimate a pooled specification where we include a full set of product-by-year effects, and a first-difference version.[13] We also estimate an analogous specification using the producer's (log) price in period *t* as the dependent variable.

The results of the TFPR level specifications are in panel A of Table 3a. Demand is positively correlated with TFPR. The estimated elasticities β_1 are positive for every product and statistically significant at the five percent level for all but two products. In the pooled specification, we estimate an average elasticity of TFPR with respect to demand of 0.064, and reject the null with a *t*-statistic of 29.9. This elasticity implies that a one standard deviation increase in plant-specific demand corresponds to an increase in TFPR of one-third of a standard deviation. The first difference specification results in panel B also reject the hypothesis of zero covariance between TFPR and demand. The pooled estimates imply that a one standard deviation increase in plant-specific demand yields an increase in TFPR of about 40 percent of a standard deviation in TFPR. (For the sake of comparison to results we describe immediately below, we also run the specification separately on the subset of our sample composed of ready-mixed concrete producers. We find similar results.)

The results using the (log) of plant-level price as the dependent variable are reported in Table 3b. The results closely mimic those for TFPR. The magnitudes of the estimated elasticities are positive and significant at the 5 percent level for seven of the eleven individual products. The pooled average elasticity of price with respect to demand is 0.059, and we reject the null with a *t*-statistic of 29.9. This elasticity implies that a one standard deviation increase in plant-specific demand corresponds to a price increase of about 30 percent of a standard deviation. The first difference results also yield a large positive and statistically significant elasticity of price with respect to plant-specific demand for the pooled sample as well as the sample restricted to ready-mixed concrete producers.

[13] We have also estimated the first difference specification industry-by-industry with year effects and pooled sample first differences with product-by-year effects and obtained very similar results. Note that sample sizes are smaller in first-difference specifications given the requirement that plants must survive and have non-imputed data in consecutive periods. We use the same inverse propensity score weight for the first differences as for the levels.

The results of these tests are especially interesting because in the prior section we highlighted that TFPQ_HK both theoretically and empirically depends strongly on demand. Under the assumptions of the HK framework, TFPR and prices should be invariant to demand, and by implication, TFPQ_HK as well. Yet when we actually measure TFPQ_HK by imposing the assumptions of the HK framework, we find that it is—contrary to the implications of the framework—correlated with price and TFPR variation. This internal inconsistency is another sign that the conditions necessary to interpret TFPR variation as reflecting distortions may not hold in the data.

We consider a second approach to testing for a relationship between TFPR and demand. This has the advantage that, in principle, one can apply it to a much wider range of data without having direct measures of prices and quantities. As such, it may be of broader applicability for researchers. It uses geographic and vertical distance measures to identify shifts in local downstream demand. We apply this methodology for the products in our dataset that are primarily sold near to where they are produced (Boxes, Bread, Concrete, and Ice), but one can imagine a much broader set of industries outside our sample to which this approach could apply.

For each of the local products, we use the detailed U.S. input/output matrix to identify the top ten downstream industries. We combine this with the Longitudinal Business Database to measure employment at the BEA Economic Area level in each downstream industry. Our downstream demand metric for each producer is the weighted average of local employment in each of the downstream demand industries (where the weights are computed using the input/output matrix). We use the log of this value in our tests.

To motivate this approach, consider ready mixed concrete. Demand for concrete is very local; almost all of it is shipped very short distances. Further, as emphasized by Syverson (2004), the construction sector accounts for 95% of the ready mixed concrete industry's revenues, but ready mixed accounts for less than 5% of construction sector's intermediate input costs. Thus (local) construction demand drives (local) ready mixed concrete outcomes and not vice versa. We extend this same logic to our other local products.[14]

As before, we consider level and first difference specifications. The former is

$$tfpr_{it} = \beta_0 + \beta_1 downdemand_{mt} + \beta_m + \eta_t + \varepsilon_{it}$$

[14] For ice and bread the top downstream industry is grocery stores. The top downstream industries for boxes are in the wholesale and retail trade sectors. In all of these industries, the share of downstream costs accounted for by the upstream industry is small, just as with concrete.

where *downdemand*$_{mt}$ is the downstream demand measure in market *m* at time *t*, η_t is a period fixed effect, and β_m is a BEA Economic Area (market) fixed effect. We estimate this specification for ready mixed concrete and a pooled estimate for all local market products. The pooled estimates include year-by-Economic Area and product-by-Economic-Area fixed effects. Standard errors are clustered by Economic Areas. Under the HK assumptions, β_1 = 0. The first difference specification uses plants that continue operations across at least two consecutive Economic Censuses. This specification is

$$\Delta tfpr_{it} = \delta_0 + \delta_1 \Delta downdemand_{mt} + \varepsilon_{it}$$

where under the HK assumptions, δ_1 = 0. SEs are again clustered by Economic Areas. We also estimate level and first difference specifications with the dependent variable based on the plant-level price.

The results of this second test using downstream demand indicators are reported in Tables 4a for TFPR and 4b for price. The magnitudes of the estimated elasticities β_1 are positive and statistically significant for price but only marginally statistically significant for the ready mixed concrete and pooled results for TFPR using the level specifications. However, the first difference specifications reject the null hypothesis of zero covariance between TFPR and demand and price and demand at a five percent level. To benchmark the magnitudes of these relationships, the first difference estimates for the pooled specification imply that a one standard deviation increase in downstream demand raises TFPR by about 35 percent of its standard deviation, the same order of magnitude of the effects we found with our other demand measures.

It seems either that the HK assumptions are violated or that distortions are positively correlated with demand. This raises an obvious question: Can one separately measure distortions and demand in the HK framework if they are in fact correlated? It would be important to do so because from a positive standpoint these firm-level primitives could have very different statistical properties (persistence, variance, etc.) as well as from a normative standpoint because clearly there are very different policy implications depending on whether firm outcomes reflect distortions or demand variation.

It turns out that, generally speaking, distortions and demand cannot be separately identified in standard production data. As we emphasize above, only with CES demand and constant marginal costs is TFPR invariant to TFPQ demand shocks in the absence of distortions. Departures from these assumptions result in TFPR being a function of fundamentals in the absence of distortions, creating an identification problem. However, with price and quantity data and assumptions about the structures of demand and technology, further progress can be made distinguishing between fundamentals and distortions. This also permits decomposing TFPR into its fundamental vs. distortion components. We explore these issues in the next section.

III. Quantifying the Effects of Departures from the HK Assumptions on Misallocation Measurement

Given the empirical findings that the HK assumptions are violated in our data, we now quantify, at least partially, the effects of departures from HK's assumptions on misallocation measurement. To implement this analysis, we require additional structure on both the demand and supply sides of the market. We consider departures from both CES demand and constant marginal costs.

A. Derivation of the Variance Decomposition of TFPR

The first step is to decompose the variance of TFPR under our more general demand and cost structures.

We can write TFPR for a producer *i* as:

$$TFPR_i \equiv P_i \cdot A_i = \frac{P_i}{MC_i} MC_i \cdot A_i = \Psi_i S_i$$

Where $\Psi_i \equiv \frac{P_i}{MC_i}$ and $S_i \equiv MC_i \cdot A_i$.

This lets us write the variance of logged $TFPR_i$ as

$$V(tfpr_i) = V(\psi_i) + V(s_i) + 2cov(\psi_i, s_i)$$

where lowercase denotes logged values. Under the HK assumptions, Ψ_i and S_i do not vary across producers, so $TFPR_i$ has a variance of zero in the absence of distortions. Here, we explore how deviations from the HK assumptions quantitatively map into TFPR variation.

To explore departures from these assumptions, we start by assuming a variable elasticity inverse demand curve that is essentially CES demand plus a quadratic term in the deviation of logged price from its average:

$$\ln[Q(P_i)] = a + b \ln P_i + d\left(\ln P_i - \overline{\ln P_i}\right)^2 + \varepsilon_i$$

This implies the demand elasticity is given by:

$$\eta = b + 2d\left(\ln P_i - \overline{\ln P_i}\right)$$

And the revenue function is:

$$R(P_i) = \widetilde{K}_i P_i^{B+1+d\ln P_i}$$

where

$$\widetilde{K}_i \equiv e^{a+d\left(\overline{\ln P_\iota}\right)^2+\varepsilon_i}$$

$$B \equiv b - 2d\,\overline{\ln P_\iota}$$

The markup is then:

$$\frac{P}{MC} = \frac{1}{1+\frac{1}{\eta}} = \frac{1}{1+\frac{1}{b+2d(p_i-\overline{p_\iota})}} = \frac{b+2d(p_i-\overline{p_\iota})}{1+b+2d(p_i-\overline{p_\iota})}$$

Taking the log of this expression yields ψ_i. Approximating it by using a first-order Taylor expansion of the first term (the logged numerator) around b and the second term (the logged denominator) around $1+b$ yields an expression for the variance of the logged markup:

$$V(\psi_i) \approx \left[\frac{2d}{b(1+b)}\right]^2 V(p_i)$$

Note that this variance is equal to zero under CES demand, where $d = 0$.

The production side of the model determines the value and variance of S_i. We consider a generalized cost function of the form

$$C(A_i, Q_i) = \left(\frac{Q_i}{A_i}\right)^{\frac{1}{\nu}} \Phi(W)$$

where ν is a scale parameter; $\nu > 1$ ($\nu < 1$) reflects economies (diseconomies) of scale. Marginal costs are then

$$MC(A_i, Q_i) = \frac{1}{\nu} Q_i^{\frac{1}{\nu}-1} A_i^{-\frac{1}{\nu}} \Phi(W)$$

Using the definition $S_i \equiv MC_i \cdot A_i$:

$$S_i = \frac{1}{\nu}\left(\frac{Q_i}{A_i}\right)^{\frac{1}{\nu}-1} \Phi(W)$$

To this point, we haven't introduced distortions into the model. Hsieh and Klenow (2009) demonstrate that revenue distortions (the τ_{Yi} in their model) are effectively shifters of the marginal cost curve.[15] We apply that logic here. For expositional convenience it is useful to specify the distortion as proportional to revenue, so we work with a distortion T_i that is equivalent to the inverse of $(1 - \tau_{Yi})$ in the HK model. In this proportional form, all distortions T_i are positive; a "tax" involves a distortion greater than one and a "subsidy" is less than one. Adding this distortion to the model implies that S_i becomes

$$S_i = \frac{1}{\nu}\left(\frac{Q_i}{A_i}\right)^{\frac{1}{\nu}-1} \Phi(W) T_i$$

Taking logs

$$s_i = \ln\frac{1}{\nu} + \ln\Phi(W) + \left(\frac{1}{\nu} - 1\right)(q_i - a_i) + \tau_i$$

Where $\tau_i \equiv \ln T_i$. The first and second terms are constants. Thus the variance of s_i is

$$V(s_i) = \left(\frac{1}{\nu} - 1\right)^2 [V(q_i) + V(a_i) - 2cov(q_i, a_i)] + V(\tau_i) + 2\left(\frac{1}{\nu} - 1\right) cov(q_i, \tau_i) - 2\left(\frac{1}{\nu} - 1\right) cov(a_i, \tau_i)$$

The final element of the variance of logged TFPR is the covariance between ψ_i and s_i. Using the expressions derived above, after some simplification we have

$$cov(\psi_i, s_i) \approx \left[\frac{2d}{b(1+b)}\right]\left(\frac{1}{\nu} - 1\right)[cov(p_i, q_i) - cov(p_i, a_i)] + \left[\frac{2d}{b(1+b)}\right] cov(p_i, \tau_i)$$

Putting the pieces together, we have that the variance of logged TFPR is

$$V(tfpr_i) \approx \left[\frac{2d}{b(1+b)}\right]^2 V(p_i) + \left(\frac{1}{\nu} - 1\right)^2 [V(q_i) + V(a_i) - 2cov(q_i, a_i)] + V(\tau_i) + 2\left(\frac{1}{\nu} - 1\right) cov(q_i, \tau_i) - 2\left(\frac{1}{\nu} - 1\right) cov(a_i, \tau_i) + 2\left[\frac{2d}{b(1+b)}\right]\left(\frac{1}{\nu} - 1\right)[cov(p_i, q_i) - cov(p_i, a_i)] + 2\left[\frac{2d}{b(1+b)}\right] cov(p_i, \tau_i)$$

Note that if $d = 0$ and $\nu = 1$ as in the HK setup, $V(tfpr_i) = V(\tau_i)$; all the variation in TFPR comes from distortions. However, if either $d \neq 0$ or $\nu \neq 1$ then TFPR will exhibit dispersion even in the absence of distortions. That is, the first, second, and fifth terms will be non-zero even if all $\tau_i = 0$. The first term will be non-zero when $d \neq 0$, the second term non-zero when $\nu \neq 1$,

[15] See their equation (6). Note that because they assume constant returns, TFPQ and the distortion term share a common unit exponent in their marginal cost expression. Here, however, because we allow non-constant returns, the exponents will differ.

and the fifth term non-zero when both $d \neq 0$ and $\nu \neq 1$. Both TFPQ and demand shocks contribute to each of these terms because the variance of prices, quantities, and the relevant covariances are a function of both.

To calculate the variance decomposition of TFPR, we need to be able to measure distortions τ_i. Our model yields an expression for the distortion. Its components can be either measured directly or estimated using our data. We derive that expression here.

Under our assumed demand, cost, and distortion structures, a firm's profit is:

$$\frac{R(P_i)}{T_i} - C\big(Q(P_i)\big) = \frac{1}{T_i}\widetilde{K}_i P_i{}^{B+1+d\ln P_i} - \left(\frac{\widetilde{K}_i P_i{}^{B+d\ln P_i}}{A_i}\right)^{\frac{1}{\nu}} \Phi(W)$$

The first order condition for a producer's price is given by:[16]

$$\frac{1}{T_i}\widetilde{K}_i(B+1+d\ln P_i)\big(P_i{}^{B+d\ln P_i}\big)\left(\frac{d}{P_i}\right)$$
$$= \frac{1}{\nu}A_i{}^{-\frac{1}{\nu}}\Phi(W)\big(\widetilde{K}_i P_i{}^{B+d\ln P_i}\big)^{\frac{1}{\nu}-1}\left(\widetilde{K}_i(B+d\ln P_i)\big(P_i{}^{B-1+d\ln P_i}\big)\left(\frac{d}{P_i}\right)\right)$$

Using our generalized revenue and cost functions, this can be used to find an expression for a producer's distortion:

$$T_i = \frac{\nu R(P_i)}{C(A,Q)}\left(\frac{B+1+d\ln P_i}{B+d\ln P_i}\right)$$

We observe revenues $R(P_i)$, total costs $C(A,Q)$, and prices directly in the data. So we just need the parameters of the demand function B and d as well as the scale elasticity ν to measure distortions.[17] These in hand, we can then compute the relevant variances and covariances to implement the TFPR variance decomposition empirically.

In the empirical analysis below we summarize this decomposition into three broad categories: the contribution of fundamentals given by the first, second, and fifth terms; the

[16] We assume an individual producer is small enough so that its price choice doesn't affect the average level $\overline{\ln P_i}$.

[17] It is easily seen that this expression nests the standard HK expressions for distortions. Under isoelastic demand (d = 0) and constant marginal costs c, the above implies that:

$$TFPR_i = P_i A_i = \left(\frac{b}{1+b}\right) cT_i$$

This again makes apparent that under the HK assumptions, variation in TFPR only depends on distortions. However, our derivation above makes clear that this property does not hold under the more general demand and production functions we consider.

contribution of distortions given by the third term; and the contribution of terms involving both fundamentals and distortions given by the fourth and sixth terms.

B. Empirical Implementation of the Variance Decomposition

We require estimates of the demand function and the scale elasticity to implement the decomposition. For the estimation of the quadratic demand function, we follow Foster et al. (2008, 2016), using TFPQ as an instrument for the main effect and squared TFPQ (deviated from its product by year mean) as an instrument for the squared term. For the estimation of the scale elasticity, we estimate the production function applying Wooldridge (2009) through two approaches. In one, we estimate returns to scale directly by measuring the elasticity of output to a composite input that is a cost-share-weighted sum of the individual logged inputs (labor, capital, materials, and energy). In the second approach, we estimate each factor elasticity separately and sum them. We find very similar estimates of returns to scale under both approaches and report the results using the composite input.

We estimate these specifications using the pooled data while controlling for product-by-year effects. This yields estimates of the pooled demand and return to scale parameters. We also estimate these parameters for selected individual products. For this purpose we restrict our attention to concrete and boxes, the two products with the largest sample size. The reason is that the Wooldridge (2009) method requires using lagged instruments in the one step GMM procedure. Moreover, the proxy methods use high order polynomials (we use cubics), and these methods are more reliable with larger samples (see Foster et al. (2017)). Concrete and boxes have sufficient number of observations to allow product-by-product implementation.

Panel A of Table 5 reports the parameter estimates. In the pooled sample, the average elasticity of demand is -2.87, but the quadratic term is large and highly significant implying considerable plant-specific variation in markups. We obtain an estimate of returns to scale statistically equal to one. For concrete, consistent with our prior work we find a larger-in-magnitude average demand elasticity, -4.18, and again a significant quadratic term. We also cannot reject constant returns to scale for this industry. For boxes, we estimate an average demand elasticity of -2.41 and a large quadratic term. Here, estimated returns to scale are 1.3, with enough precision to reject constant returns.[18]

Before turning to the decomposition, panel B of Table 5 reports key correlations. TFPR and the measure of distortions are highly correlated, but the correlation is far from one. This is especially true for boxes, where the departures from HK assumptions are the most apparent. The correlations between the measure of distortions and fundamentals (TFPQ and demand shocks) are also positive, but tend to be less positive than the correlations between TFPR and these fundamentals. This latter pattern is especially evident in the correlation with TFPQ.

Our parameter estimates imply that for the pooled and concrete samples, the second, fourth and fifth terms are zero in the above decomposition given that we cannot reject constant returns to scale. (We don't use the 1.02 estimate for concrete because we cannot reject that it equals one.) For box makers, the estimate of returns to scale of 1.3 yields non-zero contributions of those terms.

The decomposition results show that fundamentals account for an important fraction of TFPR variation that is independent of distortions. In the pooled results, about 20 percent of the variation in TFPR is accounted for by variance in fundamentals. It is roughly 140 percent for boxes. The variance of our distortion measure is also an important contributor to TFPR, and together with fundamentals accounts for more than 100 percent of the variance. This pattern holds because the terms that involve covariances between distortions and other variables have a negative contribution. For example, the sixth term in the decomposition is negative because $\frac{2d}{b(1+b)}$ is negative and the correlation between prices and distortions is positive

Our interpretation of this exercise is that we have explained some of the variation of TFPR as purely reflecting variation in demand and cost fundamentals by allowing for quadratic demand and non-constant marginal costs. Distortions do still have an explanatory role once these fundamentals are accounted for. However, it is important to note that just as with the HK framework, our distortion measure is a residual. Now, rather than being TFPR itself, it is the part of TFPR that we cannot account for with our more flexible demand and cost structures. Remaining departures of the data from *our* augmented framework would be labeled distortions even if they were not. To this point, for boxes, where we capture more of the variation in TFPR,

[18] Using a very different (Klette and Griliches (1996)) approach, Foster et al. (2017) find more extensive evidence of mild increasing returns—with estimates of returns to scale that average 1.09—on a much larger sample of industries.

the correlation between TFPR and the residual measure of distortions is substantially lower. Thus it remains an open question what our measured distortions represent. We explore this issue next by looking at their relationship to business survival.

C. *The Relationship between Measured Distortions and Survival*

Table 6 reports estimates of the marginal effect of various measures and combinations of measures on the probability of exit. Consistent with the literature and our results above in Table 2, businesses with higher TFPR, higher TFPQ, and higher demand are less likely to exit.

Interestingly, we also find that businesses with higher measured distortions are less likely to exit. Thus the results in the columns 1 and 4 of the table indicate that measured distortions—whether under the assumptions of the HK model (that is, measured as TFPR) or after accounting for departures from the HK framework in our analysis—are negatively associated with exit. More distorted producers (those facing a higher "tax") are more likely to survive. This seems an odd empirical property of distortions given the concept that underlies them.

However, the last two columns of the table show that once we control for supply and demand fundamentals, plants with higher measured distortions (again, whether using our measure or the HK-based TFPR) are in fact *less* likely to survive. This suggests measured distortions do include information about something that is a true distortion, but that this component of the measure is empirically swamped by other sources of variation that are instead associated with (positive) fundamentals about producer profitability.

These results are consistent with the identification problem we mentioned above. In the end, empirical distortion measures are a residual. Their separate identification from fundamentals exists only to the extent that one believes the modeled structure of producer demand and costs. Measured "distortions" may still embody elements of producers' idiosyncratic demand or costs that are—contrary to the concept of a distortion that acts as an implicit tax—"good news" about the producer's survival prospects. Unmodeled idiosyncratic demand and cost conditions would be misinterpreted as misallocation. We show that one can make progress on reducing the extent to which this confound occurs by using more flexible modeling structures, but we found in our sample that enough confounds remain for the "true" distortion to still be partially hidden. Moreover, we were able to leverage price and quantity data that most researchers in this literature do not have access to. Accounting for model misspecification without such data is a considerably more difficult task, raising the likelihood in more general settings that misallocation measures will confound distortions and other components of idiosyncratic profitability.

D. Other Sources of Departures from HK Assumptions

In this paper we have focused on the role of model misspecification in accounting for reasons why misallocation measures may not simply reflect wedge-like distortions. However, there are several additional reasons why revenue productivity might vary across firms in the absence of distortions. These include differences in factor prices (Katayama, Lu, and Tybout (2009)), factor quality, heterogeneity in factor demand and elasticities, adjustment costs (Asker, Collard-Wexler, and De Loecker (2014) and Decker et. al. (2017) offer extensive analysis of the role of adjustment costs) and measurement error (Bils, Klenow and Ruane (2017)).

It is beyond the scope of this paper to consider all of these alternatives, but they provide additional reasons for applying considerable caution when measuring misallocation using revenue productivity dispersion. We also mention these because we think our analysis provides guidance that can be potentially used to help differentiate amongst them. Our findings highlight the following properties observed when price and quantity data are available. First, TFPQ and TFPR are strongly positively correlated. Second, TFPR is strongly positively correlated with producers' idiosyncratic demand levels. Third, the elasticity of prices with respect to TFPQ is less than one. Fourth, survival is greater for plants with higher TFPR, higher TFPQ, and higher demand. Fifth, interestingly, when all three of these measures are considered jointly, plants with higher TFPQ and demand are more likely to survive holding the other factors constant, but plants with higher TFPR are less likely to survive holding TFPQ and demand constant. Thus, researchers should take several moments into account when evaluating models that account for TFPR dispersion.[19]

Of the alternative explanations for TFPR variation above, one that we regard as especially relevant and promising for being consistent with our evidence is factor adjustment

[19] Foster et al. (2017) include a complementary analysis with related but distinct findings. They contrast and compare TFPR to residuals from revenue function estimation, highlighting that the latter are conceptually different from TFPR. Under CES demand, the revenue function residuals are a function of fundamentals, TFPQ and demand shocks. The reason is that the estimated parameters of the revenue function are revenue elasticities reflecting both output elasticities and the demand elasticity. They find that TFPR and revenue function residuals are highly correlated, exhibit similar dispersion and are both positively related to survival. This provides a distinct set of moments that should be taken into account in modeling TFPR dispersion.

costs. A firm with a positive realization of TFPQ wants to become larger. In a frictionless environment the firm increases factors to the point where marginal revenue products equal the input factor costs. Output rises and price falls. Under the HK assumptions, price falls just enough to counteract the increase in TFPQ. If there are adjustment frictions, however, the increase in inputs and output will be smaller, making the decline in prices smaller too. Accordingly, the positive TFPQ realization will result in an increase in TFPR. Putting the pieces together, TFPR will be positively correlated with TFPQ, prices will have a less than unit elastic response to TFPQ, and—given the positive correlation between TFPR and fundamentals—higher TFPR firms will be more likely to survive. In short, adjustment frictions have implications that match many of the core findings of our analysis.

IV. **Concluding Remarks**

Measuring misallocation—identifying idiosyncratic distortions that adversely impact the allocation of resources—is a first order issue. Our analysis highlights difficult identification challenges for measuring distortions. In particular, we view our paper as sounding a note of caution about using differences across producers' measured revenue productivity (TFPR) levels to measure distortions. The stringent assumptions of the Hsieh and Klenow (2009) framework that enables such identification typically do not hold in the US data where price and quantity data are available, and other evidence suggests this may be a more general issue.

We find that there is incomplete pass-through of TFPQ in plant-level prices, one of the implications of the stringent assumptions of the HK framework. Perhaps as a result of this departure from the framework's assumptions, TFPQ measures derived indirectly using the framework are only weakly correlated with and have much more dispersion than directly measured TFPQ. Moreover, the indirect measures of TFPQ are inversely related to firm survival (in contrast to the direct measures), inconsistent with economic theory.

To quantitatively account for these patterns, we augment the HK framework to allow for departures from CES demand and constant marginal costs. We find evidence of such departures in our data and measure the extent to which they result in non-distortionary demand and cost fundamentals create dispersion in TFPR. We also compute the residual measure of distortions that emerges from our model. We find it is highly correlated with TFPR and demand and cost fundamentals. It is also positively related to survival, again at odds with the concept of misallocation-inducing distortions. One core message of our analysis is that our residual measure of distortions (as well as TFPR, the distortion measure in the HK framework) is highly correlated with directly measured fundamentals and has similar relationships with survival. These findings taken together sound a note of caution in using TFP, or even our residual measure, as a measure of distortions.

While much of the message of our paper is to sound a note of caution, one of our findings suggests there is interesting information captured by TFPR (and our residual measure) once one controls for fundamentals. Specifically, we find that after controlling for TFPQ and demand shocks, high TFPR and high residual measures of distortion plants are more likely to exit. Thus, if independent information on fundamentals can be measured (feasible with price and quantity data but a challenge in their absence) then one might be able to make progress at isolating a sharper measures of true distortions in the data.

References

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

Baily, Martin, Eric Bartelsman and John Haltiwanger. 2001. "Labor Productivity: Structural Change and Cyclical Dynamics," *Review of Economics and Statistics*, 83(3): 420–433.

Bonnet, Céline, Pierre Dubois, Sofia B. Villas Boas, and Daniel Klapper. 2013. "Empirical Evidence on the Role of Nonlinear Wholesale Pricing and Vertical Restraints on Cost Pass-Through." *Review of Economics and Statistics*, 95(2): 500–15.

Campa, José Manuel and Linda S. Goldberg. 2005. "Exchange Rate Pass-Through into Import Prices." *Review of Economics and Statistics*, 87(4): 679–90.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2017. "Changing Business Dynamics and Productivity: Shocks vs. Responsiveness." Working Paper.

De Loecker, Jan and Frederic Warzynski. 2012. "Markups and Firm-level Export Status." *American Economic Review*, 102(6), 2437-71.

Fabra, Natalia and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *American Economic Review*, 104(9): 2872-99.

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

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

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

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2016. "Firm-Level Dispersion in Productivity: Is the Devil in the Details?" *American Economic Review Papers and Proceedings*, May.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2017. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights." NBER Working Paper No. 23666.

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker. 2016. "Energy Prices, Pass-Through, and Incidence in U.S. Manufacturing." Working paper.

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

Haltiwanger, John. Forthcoming. Firm Dynamics and Productivity: TFPQ, TFPR and Demand Side Factors." *Economia*.

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

Katayama, Hajime, Shihua Lu, and James R. Tybout. 2009. “Firm-Level Productivity Studies: Illusions and a Solution.” *International Journal of Industrial Organization*, 27(3): 403–13.

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): 343–361.

Kulick, Robert. 2016. “Horizontal Mergers, Prices and Productivity,” mimeo.

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

Shea, John. 1993a. “The Input-Output Approach to Instrument Selection.” *Journal of Business & Economic Statistics*, 11(2): 145-55.

Shea, John. 1993b. “Do Supply Curves Slope Up?” *Quarterly Journal of Economics*, 108(1): 1-32.

White, T. Kirk, Jerome P. Reiter, and Amil Petrin. 2014. “Plant-level Productivity and Imputation of Missing Data in U.S. Census Manufacturing Data.” Mimeo, Center for Economic Studies.

Wooldridge, Jeffrey. 2009. “On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables.” *Economics Letters*, 104(3): 112-114.

Figure 1. Effect of a Change in TFPQ in the Hsieh-Klenow Framework

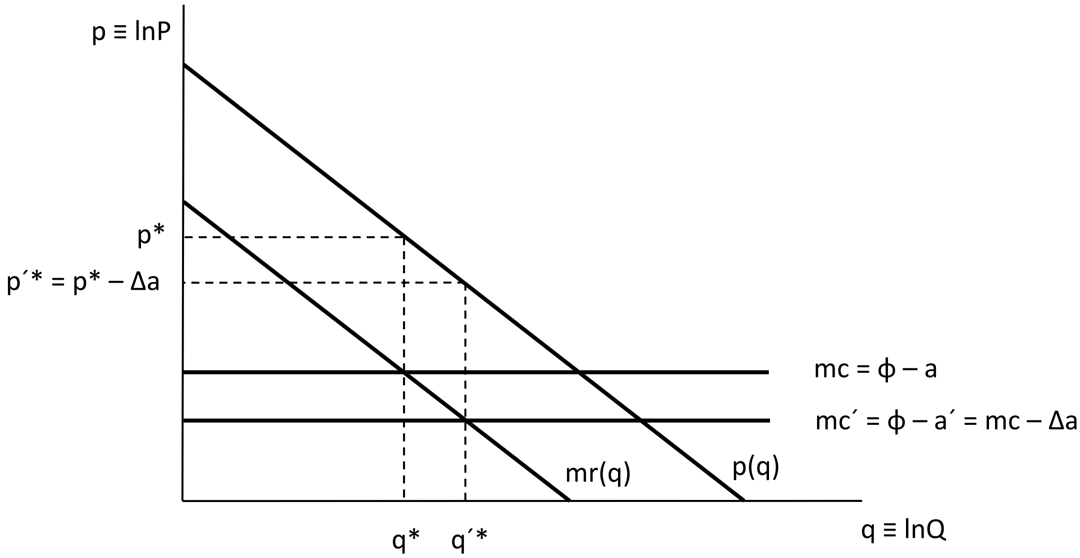


Figure 2. Effect of a Change in TFPQ when the Marginal Cost Curve Is Not Horizontal

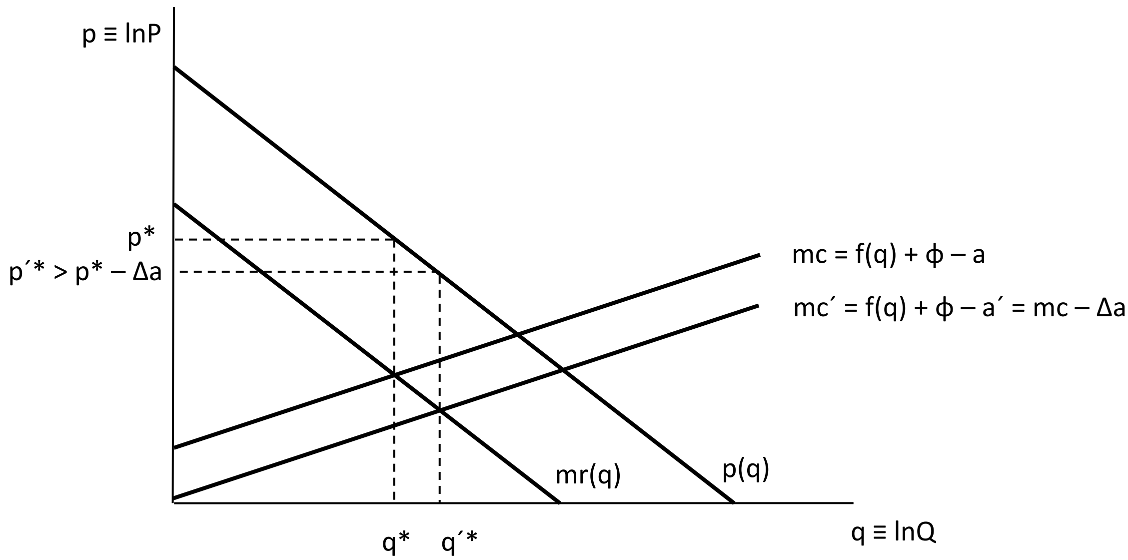


Figure 3. Demand Shifts Do Not Change TFPR in HK

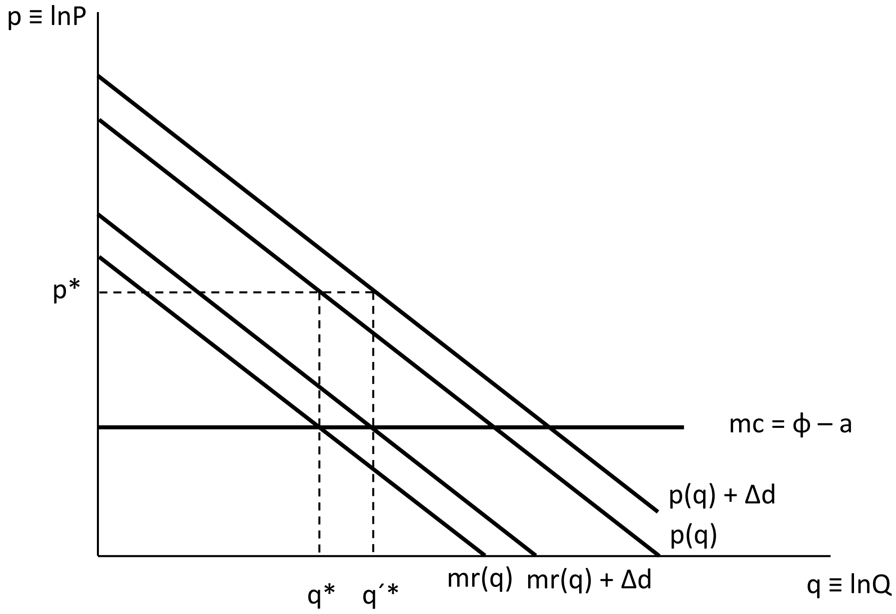
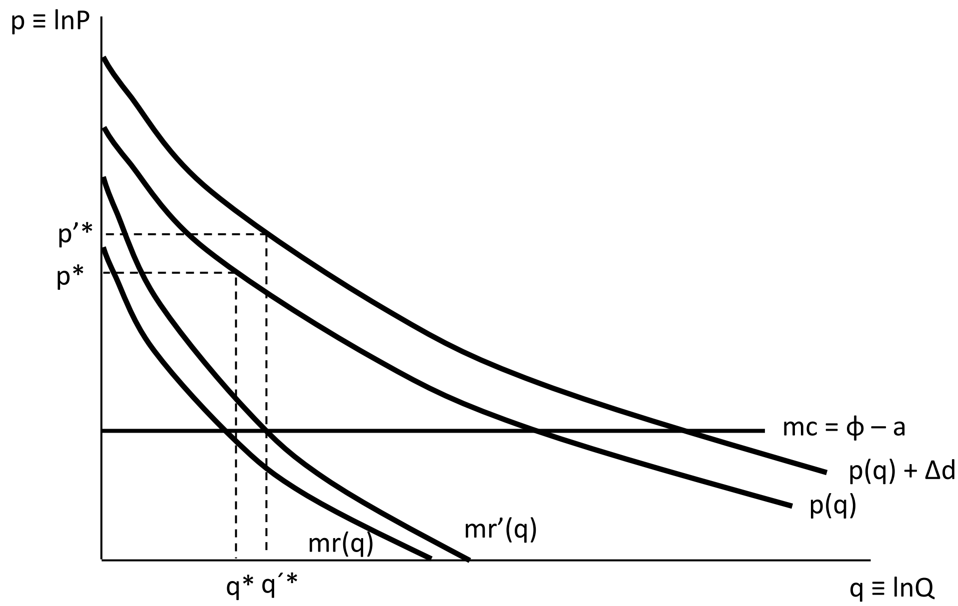


Figure 4. Demand Shifts Change TFPR If HK Assumptions Do Not Hold

A. Non-Isoelastic Demand but Constant Marginal Costs



B. Isoelastic Demand but Non-Constant Marginal Costs

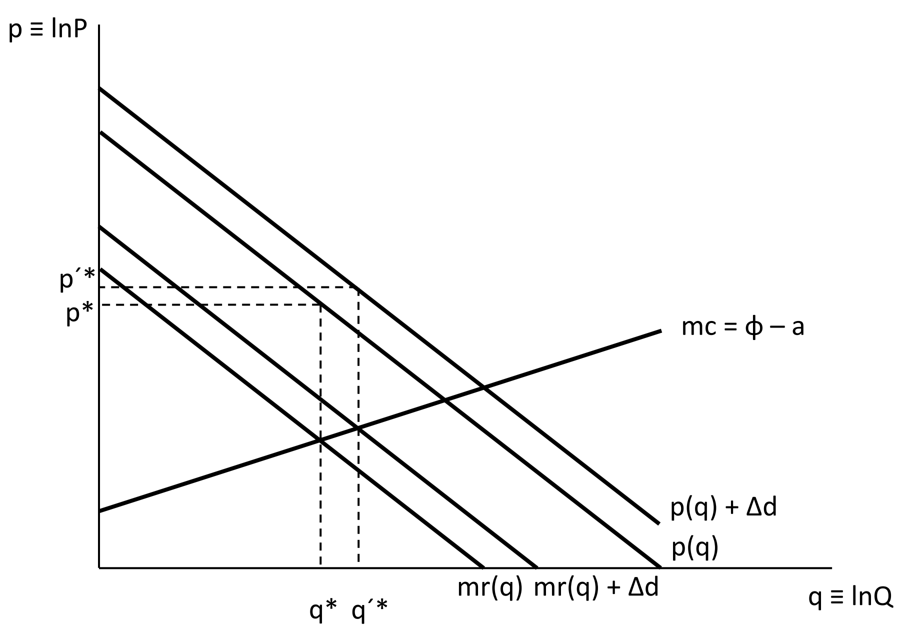


Table 1. Elasticity of Plant-Level log(Price) to log(TFPQ)

Product	Point Estimate	Std. Error	t-stat for H_0: $\alpha_1 = -1$
Boxes	-0.825	0.013	-13.4
Bread	-0.521	0.031	-15.6
Carbon Black	-0.691	0.071	-4.4
Coffee	-0.527	0.038	-12.5
Concrete	-0.265	0.008	-91.9
Flooring	-0.724	0.064	-4.3
Gasoline	-0.251	0.024	-31.3
Block Ice	-0.569	0.067	-6.4
Processed Ice	-0.521	0.041	-11.8
Plywood	-0.862	0.020	-6.9
Sugar	-0.177	0.035	-23.5
Pooled, OLS	-0.450	0.006	-86.4
Pooled, IV (Innovation to TFPQ)	-0.420	0.017	-35.1
Pooled, IV (Lagged TFPQ)	-0.537	0.043	-10.7

Notes: The total sample (pooled) is approximately 9500 observations. All specifications use inverse propensity score weights to account for selection in using only non-imputed physical product data. By product estimates include year effects. Pooled specifications include product by year effects. Differences in reported t-statistics and ratio of reported point estimates and standard errors subject to rounding error.

Table 2. Selection on Alternative TFP Measures and Demand

Specification:	[1]	[2]	[3]	[4]
TFPR	-0.036* (0.021)			
TFPQ		-0.035* (0.018)		
Demand Shock			-0.056*** (0.005)	
TFPQ_HK				-0.008*** (0.002)

Note: These results are from various probits of plant exit by the next census (shown by column) on plant-level (logged) productivity measures as well as a full set of product-year fixed effects. The sample is the pooled sample of approximately 9200 observations. Standard errors, clustered by plant, are in parentheses.

Table 3a. Elasticity of Plant-Level ln(TFPR) to Plant-Level ln(Demand)

A. Levels:

Product	Point Estimate	Std. Error	t-stat for H_0: $\beta_1 = 0$
Boxes	0.029	0.003	8.8
Bread	0.118	0.010	12.4
Carbon Black	0.087	0.045	1.9
Coffee	0.074	0.008	9.3
Concrete	0.068	0.003	24.2
Flooring	0.069	0.028	2.4
Gasoline	0.004	0.005	0.7
Block Ice	0.195	0.060	3.3
Processed Ice	0.098	0.030	3.2
Plywood	0.008	0.015	0.5
Sugar	0.085	0.031	2.8
Pooled, All Products	0.064	0.002	29.9

B. First Difference Specification for Continuing Plants:

Product	Point Estimate	Std. Error	t-stat for H_0: $\delta_1 = 0$
Concrete	0.135	0.007	20.7
Pooled, All Products	0.133	0.005	25.8

Notes: The total sample (pooled) is approximately 9500 observations. All specifications use inverse propensity score weights to account for selection in using only non-imputed physical product data. For the level specifications, by product estimates include year effects and pooled specification includes product by year effects. For the first difference specification, pooled specification includes product effects. Differences in reported *t*-statistics and ratio of reported point estimates and standard errors subject to rounding error.

Table 3b. Elasticity of Plant-Level ln(Price) to Plant-Level ln(Demand)

A. Levels:

Product	Point Estimate	Std. Error	t-stat for H0: $\beta_1 = 0$
Boxes	0.028	0.006	4.9
Bread	0.118	0.010	11.8
Carbon Black	0.054	0.059	0.9
Coffee	0.074	0.008	8.8
Concrete	0.061	0.002	32.4
Flooring	0.068	0.044	1.6
Gasoline	0.004	0.003	1.1
Block Ice	0.192	0.069	2.8
Processed Ice	0.113	0.031	3.6
Plywood	-0.001	0.043	0.0
Sugar	0.071	0.015	4.9
Pooled, All Products	0.059	0.002	29.9

B. First Difference Specification for Continuing Plants:

Product	Point Estimate	Std. Error	t-stat for H0: $\delta_1 = 0$
Concrete	0.133	0.003	40.8
Pooled, All Products	0.159	0.003	46.2

Notes: The total sample (pooled) is approximately 9500 observations. All specifications use inverse propensity score weights to account for selection in using only non-imputed physical product data. For the level specifications, by product estimates include year effects and pooled specification includes product by year effects. For the first difference specification, pooled specification includes product effects. Differences in reported *t*-statistics and ratio of reported point estimates and standard errors subject to rounding error.

Table 4a. Elasticity of Plant-Level ln(TFPR) to Downstream Demand

A. Level:

Product	Point Estimate	Std. Error	*t*-stat for H_0: $\beta_1 = 0$
Concrete	0.046	0.025	1.82
Pooled, Local Products	0.042	0.024	1.74

B. First Difference Specification for Continuing Plants:

Product	Point Estimate	Std. Error	*t*-stat for H_0: $\delta_1 = 0$
Concrete	0.127	0.052	2.42
Pooled, Local Products	0.115	0.050	2.33

Table 4b. Elasticity of Plant-Level ln(Price) to Downstream Demand

C. Level:

Product	Point Estimate	Std. Error	*t*-stat for H_0: $\beta_1 = 0$
Concrete	0.075	0.022	3.42
Pooled, Local Products	0.076	0.022	3.51

D. First Difference Specification for Continuing Plants:

Product	Point Estimate	Std. Error	*t*-stat for H_0: $\delta_1 = 0$
Concrete	0.108	0.032	3.42
Pooled, Local Products	0.107	0.029	3.72

Notes: The total sample (pooled for local products) is approximately 8000 observations. All specifications use inverse propensity score weights to account for selection in using only non-imputed physical product data. For the level specifications, by product estimates include year effects and economic area effects, and pooled specification includes product, year and economic area effects. For the first difference specification, pooled specification includes product effects. Differences in reported t-statistics and ratio of reported point estimates and standard errors subject to rounding error.

Table 5: Distortion Estimates with Quadratic Demand and Non-constant Marginal Costs

Panel A: Parameter Estimates

Statistic	Pooled	Concrete	Boxes
Estimate of b	-2.87	-4.18	-2.41
	(0.09)	(0.24)	(0.15)
Estimate of d	-0.96	-1.89	-3.25
	(0.15)	(1.07)	(0.39)
Estimate of υ	1.00	1.02	1.30
	(0.02)	(0.02)	(0.07)

Panel B: Correlations between Distortions and Fundamentals

Statistic	Pooled	Concrete	Boxes
Corr(p_i, τ_i)	0.39	0.38	0.67
Corr($tfpr_i, \tau_i$)	0.83	0.84	0.51
Corr(a_i, τ_i)	0.39	0.55	0.04
Corr(ε_i, τ_i)	0.25	0.27	0.13
Corr($tfpr_i, a_i$)	0.66	0.78	0.73
Corr($tfpr_i, \varepsilon_i$)	0.27	0.31	-0.62

Panel C: Variance Decomposition

Fraction of Variance of TFPR from:	Pooled	Concrete	Boxes
Fundamentals	0.21	0.05	1.39
Distortions	1.28	1.13	0.20
Covariance of fundamentals and distortions	-0.50	-0.18	-0.59

Notes: Standard errors in parentheses in panel A. Pooled results control for product by year effects. In panel C, “Fundamentals” include the first, second, and fifth terms of the decomposition. “Distortions” reflect the variance of the third term, and the “Covariance” terms are the fourth and sixth terms of the decomposition.

Table 6. Selection on TFP Measures, Demand, and Distortions

Specification:	[1]	[2]	[3]	[4]	[5]	[6]
TFPR	-0.036*					0.138***
	(0.021)					(0.030)
TFPQ		-0.035*			-0.041**	-0.107***
		(0.018)			(0.018)	(0.024)
Demand Shock			-0.056***		-0.059***	-0.064***
			(0.005)		(0.005)	(0.005)
Distortions (τ)				-0.039**	0.049***	
				(0.017)	(0.018)	

Note: These results are from various probits of plant exit by the next census (shown by column) on plant-level productivity measures as well as a full set of product-year fixed effects. The sample is the pooled sample of approximately 9200 observations. Standard errors, clustered by plant, are in parentheses.