Misallocation Measures:

Glowing Like the Metal on the Edge of a Knife

John Haltiwanger, University of Maryland and NBER

Robert Kulick, University of Maryland

Chad Syverson, U. of Chicago Booth School of Business and NBER

Misallocation

Misallocation: the covariance between size and productivity viewed through the lens of optimality

Are producers the “right” size?

Empirical cornerstone of misallocation literature: Hsieh and Klenow (2009)

Huge literature:

- HK is the most cited *QJE* article (2nd most cited top-5 journal article) published since 2008
- HK method has been widely applied in empirical settings

Hsieh-Klenow Method

Appeal of HK is obvious

- Model-based empirical method
    - Interpretable metrics (wedges)
    - Facilitates counterfactual exercises (e.g., how much would aggregate TFP rise in absence of misallocation)

- Lots of empirical power and flexibility
    - Works with “standard” production data (e.g. just revenues, no P-Q breakdown)
    - Yields producer-year panel of 2-D distortions/wedges/frictions that can be examined further and related to many other objects

Hsieh-Klenow Method

Model-based empirics depend on the model's ability to explain the data

Key implication of HK model—in absence of misallocation, revenue-based TFP (TFPR, revenue per unit input) of all producers should be the same—holds in (essentially) one case

Requires assumed functional forms on both supply and demand side of market to hold

Our Study: Questions We Want to Answer

What happens if the assumptions do not hold? Is it random measurement error or something more systematic?

Do the assumptions actually hold in the data?

What quantitative deviations should we expect in interpretation of identified distortions if assumptions don’t exactly hold?

Our Study: Answers (So Far)

What happens if the assumptions do not hold? Is it random measurement error or something more systematic? More systematic; “good” shocks to fundamentals often imply larger measured distortions; also, model seems to “miss” on one particular side of the equation

Do the assumptions actually hold in the data? Not in our (small but unusually detailed) sample; some hints they don’t more generally

What quantitative deviations should we expect in interpretation of identified distortions if assumptions don’t exactly hold? Work in progress

Hsieh-Klenow Method Recap

Industry a continuum of monopolistically competitive firms, indexed by i

Each makes product variety with its own TFPQ level A_i

Dixit-Stiglitz demand for industry's varieties

Yields residual demand curve $P_i = Q_i^{-\sigma}$

Firm production function is Cobb Douglas CRTS: $Q_i = A_i L_i^{\alpha} K_i^{1-\alpha}$.

Firms choose a quantity (equivalently, price) to maximize profit:

$$\pi_i = (1 - \tau_{Yi}) P_i Q_i - W L_i - (1 + \tau_{Ki}) R K_i$$

Nonstandard elements here are two distortions τ_{Yi} and τ_{Ki}

Hsieh-Klenow Method Recap

Under model's assumptions, TFPR is proportional to weighted geometric average of MRPL and MRPK

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

Thus only firm-level variables that shift $TFPR_i$ are distortions

If there are differences in TFPR across producers, HK framework interprets them as reflecting distortions

Hsieh-Klenow Method Recap

There are other things that would create TFPR dispersion in data:

- Different factor prices
- Different technologies (α)
- Adjustment costs (Asker, Collard-Wexler, De Loecker (2014))
- Measurement error (Bils, Klenow, and Ruane (2017))

We set these aside

We explore possible TFPR dispersion from model misspecification: deviations in assumptions about the structure of demand or technology (even when homogeneous)

Underneath the HK Result

TFPR definition: $TFPR_i \equiv P_i \cdot TFPQ_i = P_i \cdot A_i$

What makes $TFPR_i$ invariant across heterogeneous-productivity firms?

The elasticity of P_i w.r.t. A_i needs to be -1

In HK model, the condition holds only in absence of distortions:

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

Underneath the HK Result

Digging deeper into TFPR invariance condition (suppressing firm index):

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

Notes on this condition:

- $\varepsilon_{P,MC}$ depends on residual demand curve
- $\varepsilon_{MC,A}$ depends on marginal cost curve (production function)
- Demand and supply not completely independent, however
  - Must hold at the profit-maximizing price and quantity, which depend on intersection of MR & MC
- Must hold at *any* profit-maximizing quantity industry firms might operate at given the A_i distribution

The Demand-Side Assumption

When $\varepsilon_{P,MC} = 1$, price is constant multiplicative markup of MC: $P = \mu \cdot MC$

This requires isoelastic residual demand, $Q = DP^{-\sigma}$

Many standard functional forms for demand imply $\varepsilon_{P,MC} < 1$

E.g., linear demand: $\varepsilon_{P,MC} = (1/2)(MC/P)$
For any $P \geq MC$, $\varepsilon_{P,MC} \leq ½$

Condition $\varepsilon_{P,MC} < 1$ implies positive correlation between TFPQ and TFPR, as is typically found in data

The Supply-Side Assumption

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

Total change in *MC* induced by change in *A* depends on

- Direct negative effect of TFPQ on costs—shift in MC curve
- Effect of TFPQ on optimal quantity—movement along MC curve

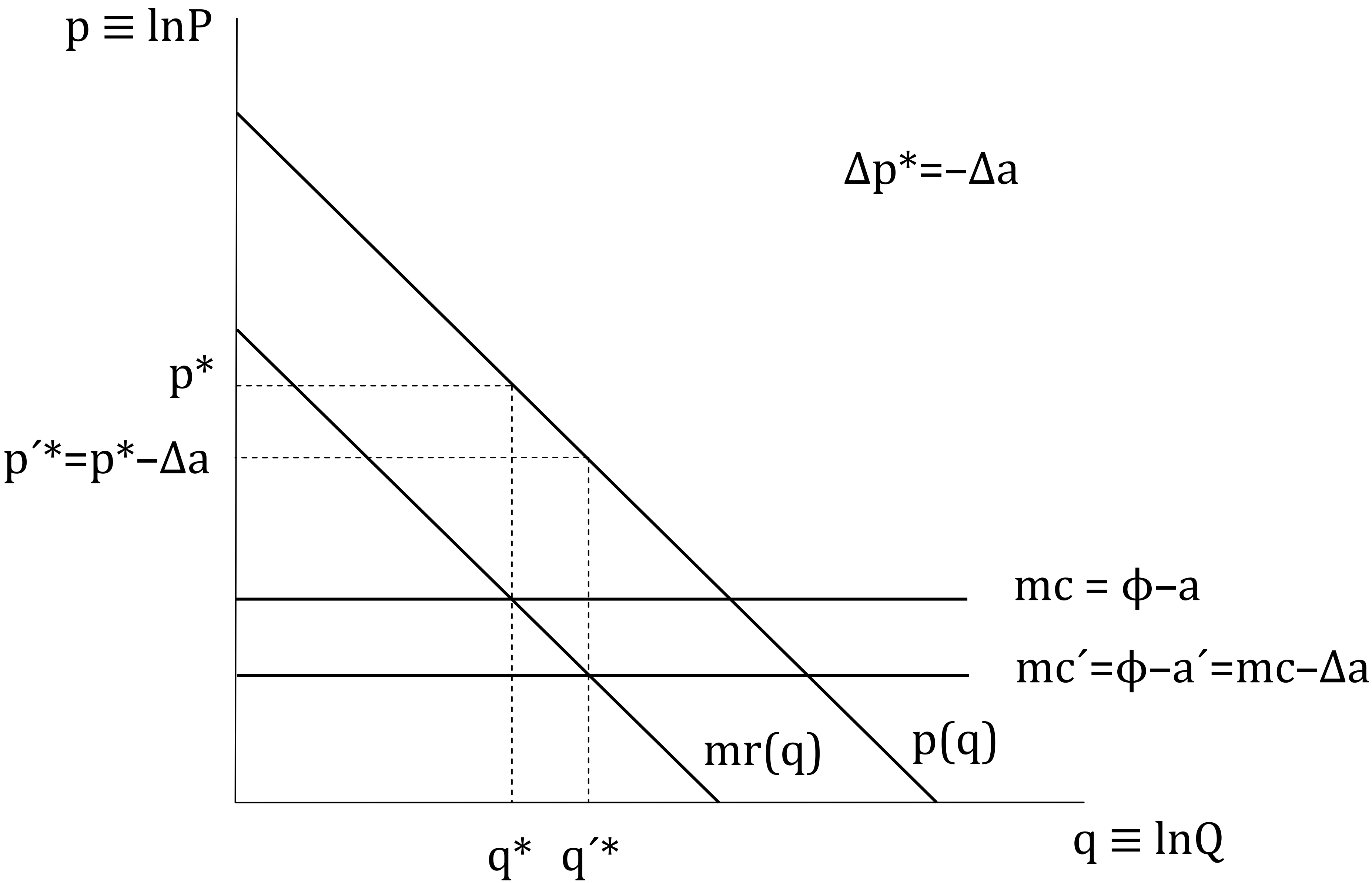
Simplest way to get $\varepsilon_{MC,A} = -1$ is flat MC curve with marginal costs negative unit elastic in TFPQ

Example: Cost Function for Cobb-Douglas PF, $Q = AL^{\alpha}K^{\beta}$

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

Conforms to HK assumption only if $\alpha + \beta = 1$; i.e., PF must have CRTS

Uniqueness of the HK Assumptions: Graphics



What Do We Know about TFPR vs. TFPQ Empirically?

For studies with data with both P and Q at micro level that allow both TFPR and TFPQ to be directly calculated:

- TFPR has slightly lower dispersion than TFPQ
- Prices are declining in TFPQ
- TFPR is positively correlated with TFPQ and idiosyncratic demand
- High TFPR firms are more likely to grow, export, and survive

In much of the literature, TFPR is treated as a proxy for TFPQ. HK argues this is misleading.

Testing the Assumptions (I)

We test whether prices are negative unit elastic in TFPQ

Requires prices

Backing out unobservable P and Q information from revenue data require assumptions; any test would be joint test of HK assumptions and the assumptions of these techniques

But we constructed dataset in our earlier work (Foster, Haltiwanger, and Syverson, 2008, 2016) that has producer-level Q and P

Extract of 1977, 1982, 1987, 1992, and 1997 U.S. Census of Manufactures with P and Q data for 11 products

Testing the Assumptions (I)

Test industry-by-industry and on pooled sample

Specification:

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

where η_t is a time (CM) fixed effect

Under the HK assumptions, α_1 = -1

Testing the Assumption (I)

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 (Innov. to TFPQ)	-0.420	0.017	-35.1
Pooled, IV (Lagged TFPQ)	-0.537	0.043	-10.7

$$\left|\varepsilon_{P,A}\right| < 1 \Leftrightarrow \text{cov(TFPR, TFPQ)} > 0$$

Testing the Assumptions (I)

Our sample is small and non-representative

However, as noted above, when researchers have both Q and P data, always find cov(TFPR,TFPQ) > 0

Thus $|\varepsilon_{P,A}| < 1$ is likely to hold more generally

Plus there is an entire literature on pass-through, with well identified tests of how much changes in costs show up in changes in prices.

- Most of the time the results indicate incomplete pass-through
- Again, points to $|\varepsilon_{P,A}| < 1$

Testing the Assumptions (II)

HK method allows TFPQ to be backed out from revenue and input data. We can compare this measure (TFPQ_HK) to our directly measured TFPQ

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

where κ is a collection of constants that are the same across all producers

Direct measure of TFPQ:

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

Testing the Assumptions (II)

Moment	11 Products	Excluding low elasiticity
SD of ln(TFPQ_HK)	3.29	1.02
SD of ln(TFPQ)	0.28	0.28
Corr(TFPQ_HK, TFPQ)	0.09	0.29
Corr(TFPQ_HK, Price)	0.006	0.014
Corr(TFPQ, Price)	-0.59	-0.59

Very poor correspondence partly due to industries with low (close to -1) price elasticities, but even excluding those, TFPQ_HK doesn't look much like TFPQ

- TFPQ_HK much more dispersed than TFPQ
- Correlation is 0.29
- TFPQ_HK is uncorrelated with price
- TFPQ strongly negatively correlated with price

Testing the Assumptions (III)

Under the HK assumptions, demand shifts don't change TFPR, *even in the presence of distortions* (under the null that distortions are uncorrelated with demand)

Intuition:

- Flat MC curve, so demand shift doesn't change MC
  - Distortions shift MC up and down, but do not change its slope
- Isoelastic demand, so markup never changes
- Therefore demand shift doesn't change $P = \mu \cdot MC$
- Therefore demand shift doesn't change $TFPR = P \cdot TFPQ$

However, any deviation from HK causes demand shifts to change price and therefore TFPR

Testing the Assumptions (III)

Foster, Haltiwanger, and Syverson (2008) actually did a test like this (of course not recognizing it at the time)

Our producer-specific demand measure was by construction the variation in firms' quantities sold that was orthogonal to costs—just like the kind assumed under the null hypothesis

We found this demand shift was significantly and positively correlated TFPR ($\rho = 0.29$), rejecting the HK assumptions

Testing the Assumptions (III)

We also use IO matrix to identify downstream demand indicators (like in Bartelsman et al. (1998) and Syverson (2004))

We estimate the following in both levels and first differences:

$$tfpr_{it} = \beta_0 + \beta_1 Downdemand_{mt} + \delta_m + \eta_t + \varepsilon_{it}$$

Under the HK assumptions, $\beta_1 = 0$

Results for levels specification: $\hat{\beta}_1 = 0.042$ (s.e. = 0.024)
Results for first-difference specification: $\hat{\beta}_1 = 0.115$ (s.e. = 0.050)

Magnitudes: 1-SD shift in demand $\Rightarrow$ 0.25 to 0.35-SD shift in TFPR

Findings fit with the more general pattern that TFPR is highly correlated with other fundamentals such as TFPQ

Quantifying Departures from HK Assumptions

Rewrite TFPR:

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

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

Thus variance of logged $TFPR_i$ is

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

Under HK assumptions, Ψ_i and S_i don't vary across producers, so $TFPR_i$ has a variance of zero

We can specify a non-CES demand system and non-CRTS cost function and compute what portion of V($tfpr_i$) is explained by departures from HK

Conclusions (Very Tentative)

HK method powerful but power comes with tight assumptions

In our sample, direct test of key assumption—response of prices to differences in TFPQ—fails; generality implied by pass-through literature

TFPQ values backed out from HK model don't look like directly computed TFPQ

TFPR strongly correlated with demand shocks

We're working on quantitatively mapping departures from assumptions

None of this suggests misallocations aren't an important source of aggregate productivity variation

Conclusions (Very Tentative)

So, what to do?

Be mindful of potential model misspecification when working with distortion metrics (in addition to other potential issues like measurement error, adjustment costs,…)

Whenever possible, find independent measures of things that are plausible proxies for distortions and show they are correlated with model-based measures