STEG Lecture Series on Macro Development: Misallocation

Richard Rogerson

Princeton University

March 2021

Motivation

- Growth accounting exercises → aggregate differences in TFP are an important source of cross-country differences in living standards.

- Highlights the need to understand the forces that shape TFP differences.

- Two distinct channels:

1. Differences in the adoption of best practice methods and technologies.

2. Differences in the extent to which resources are allocated efficiently.

- Models that start with an aggregate production function cannot distinguish between the two.

Outline

1. Benchmark models of misallocation (i.e., models of the firm size distribution).

2. Suggested channels for TFP effects

3. Misallocation as a source of TFP differences

4. Restuccia and Rogerson (RED 2008)

5. Hsieh and Klenow (QJE 2009)

6. Issues

Three Models of Cross-Sectional Production Heterogeneity

- Hopenhayn model

- Melitz model

- Lucas span of control model

We Proceed in 3 steps:

Step 1: Static models without heterogeneity

Step 2: Static models with heterogeneity

Step 3: Dynamic models with heterogeneity

- Single household, with one unit of time, preferences $u(c)$.

- Unit of production is a plant.

  - Creating a plant requires $\bar{h}$ units of labor.

  - Once created, a plant i has production function $y_i = zf(h_i)$

- f is DRS, and DRS is important.

Melitz (Really Dixit-Stiglitz)

- Single household, with one unit of time, preferences $u(c)$.

- Unit of production is a plant that produces a single variety i.

    - Creating a plant requires $\bar{h}$ units of labor.

    - Once created, plant produces output of variety i via $y_i = zh_i$

(Note: We assume CRS, but could assume DRS)

- Varieties aggregated via:

$$y = [\int_0^N y_i^\rho di]^{1/\rho}, \, 0 < \rho < 1$$

Lucas

- Unit mass of identical households, unit of time each, preferences $u(c)$.

- Unit of production is a plant.

    - A plant requires a manager and some workers.

    - Given a manager, plant i with h workers produces $y_i = zf(h_i)$.

- f is again DRS .

- Note: Looks like Hopenhayn w/o $\bar{h}$, but need for a manager is like $\bar{h} = 1$.

Efficient Allocations

Symmetry $\rightarrow$ all plants identical $\rightarrow$ Three SP problems:

Hopenhayn:

$$\max_N u\left(Nzf\left(\frac{1-N\bar{h}}{N}\right)\right) \rightarrow \max_N Nf\left(\frac{1-N\bar{h}}{N}\right)$$

Melitz:

$$\max_N u\left(N^{1/\rho}z\frac{1-N\bar{h}}{N}\right) \rightarrow \max_N N^{1/\rho}z\frac{1-N\bar{h}}{N}$$

Lucas:

$$\max_N u\left(Nzf\left(\frac{1-N}{N}\right)\right) \rightarrow \max_N Nf\left(\frac{1-N}{N}\right)$$

Solutions for a Special Case

If $f(y) = y^{\alpha}$ we have the following solutions:

Hopenhayn:

$$N = \frac{1 - \alpha}{\bar{h}}$$

Melitz:

$$N = \frac{1 - \rho}{\bar{h}}$$

Lucas:

$$N = 1 - \alpha$$

Decentralizations

Hopenhayn: standard equilibrium notion is competitive equilibrium.

Melitz: standard equilibrium notion is monopolistic competition in market for varieties and competition elswhere. (Why?)

Lucas: standard equilibrium notion is competitive equilibrium in which manager is residual claimant on proceeds of a plant.

With inelastic labor, each decentralization generates efficient allocations.

Summary

- All three models endogenously determine the intensive and extensive margins of production–the number of producers and the size of each producer.

- But producer size distribution is degenerate.

- Previous analysis assumed that all plants had the same z.

- Robust empirical finding: large cross-section dispersion in z even within narrowly defined industries.

- We now show how all three models will generate non-degenerate size distributions if we incorporate heterogeneous z.

Ex-Ante vs Ex-Post Heterogeneity

- Two different ways that we could introduce the heterogeneity:

  - ex ante heterogeneity (z_i observed after plant is created)
  - ex-post heterogeneity (z_i observed after plant is created)

- No right or wrong; both are plausible, but the choice does matter.

- Hopenhayn and Melitz traditionally assume ex-post heterogeneity; use $\bar{h}$ units of labor to get a random draw z_i.

- Lucas SoC traditionally assumes ex ante heterogeneity, calling it heterogeneous manager ability. High z_i become managers (selection).

Hopenhayn With Ex-Post Heterogeneity

- If a plant is created, it draws z_i from a distribution with density $g(z)$.

- Draws iid across plants→law of large numbers → dist. of z across plants is $g(z)$

- Social Planner now solves:

$$\max_{N, h(z)} u\left(N \int z h(z)^\alpha g(z)\right) \text{ s.t. } N \int h(z) g(z) = 1 - N\bar{h}$$

Given N, $h(z)$ must equate MPL across plants:

$$z \alpha h(z)^{\alpha - 1} = \text{constant for all } z$$

Solving the Model

Step 1: Solve for the maximal output conditional on a unit of labor and a unit mass of plants with distribution $g(z)$. Call this $Y(1)$.

Step 2: Show that $Y(H) = H^\alpha Y(1)$.

$$\frac{h(z')}{h(z'')} = [\frac{z''}{z'}]^{1/(\alpha-1)} \rightarrow \text{each } h(z) \text{ scales w/ } H$$

Step 3: Solve for optimal N by noting that total output is $N(\frac{1-N\bar{h}}{N})^\alpha Y(1)$.

Note: solution for N is unaffected; model is like homogeneous model with $z = Y(1)$.

Hopenhayn With Ex-Post Heterogeneity II

- We split $\bar{h}$ into two parts: $\bar{h}_e$ and $\bar{h}_f$.
  - Creating a plant requires $\bar{h}_e$, $\longrightarrow$ random draw from $g(z)$
  - Operating a plant requires $\bar{h}_f$, $\longrightarrow$ creates selection margin

- Social Planner now chooses N, $h(z)$, $o(z)$ to solve:

$$\max u(N \int o(z)zh(z)^{\alpha}g(z)dz) \text{ st } N \int o(z)(h(z)+\bar{h}_f)g(z)dz = 1 - N\bar{h}_e$$

- Given N and $o(z)$, $h(z)$ must satisfy: $z\alpha h(z)^{\alpha-1} =$constant for all z with $o(z) = 1$

- $o(z) \in \{0, 1\}$ will follow a threshold rule: $o(z) = 1$ if $z \geq z^*$

Melitz, Lucas With Heterogeneity

- Other two models can be extended similarly.

- Melitz extensions are essentially identical.

- Lucas is traditionally different since it specifies distribution of managerial ability (also called span of control) as known ex ante.

Decentralizations: Hopenhayn vs. Melitz

- An important distinction arises when we move to heterogeneous production units.

- Hopenhayn and Lucas: all units produce the same identical good → a single price p.

- Melitz: each unit produces a unique variety.

  - Prices are a fixed markup over marginal costs.

  - Heterogeneous productivities → heterogeneous prices.

- This distinction will be relevant when connecting to the data in Hsieh-Klenow (QJE 2009).

Dimensions of Efficient Allocations

- If we started with an aggregate production function ($Y = F(H)$), efficiency is simply $H = 1$.

- In Hopenhayn w/o heterogeneity, efficiency specifies the number of producers created and the (degenerate) distribution of workers across them.

- In Hopenhayn w/ ex-post heterogeneity, efficiency specifies the number of producers created and the non-degenerate distribution of workers across them.

- In Hopenhayn w/ ex-ante and ex-post heterogeneity, efficiency specifies number of producers created, which of them operate, and the distribution of workers across them.

Dynamic Hopenhayn Model: Households

- Representative consumer

- Preferences

$$\sum_{t=0}^{\infty} \beta^t u(C_t, 1 - H_t)$$

- Endowed with one unit of time in each period

- Production unit is the plant

- Plant production function:

$$y_t = z_t f(k_t, h_t)$$

- z_t markov chain on $Z = \{z^1, z^2, ...z^N\}$, transition matrix Q_{ij}

- z_t observed at beginning of period before any decisions

- f displays decreasing returns to scale in (k, h) jointly

- there is a fixed cost $\bar{h}_f$ in units of labor of remaining in existence

- there is a fixed cost $\bar{h}_e$ in units of labor of entering

- new plants draw their initial value of z from a distribution G ($= [G_1....G_N]$) on Z. Draws are iid across plants.

Equilibrium

- Formulate equilibrium recursively.

- Household owns the capital and rents it to plants.

- **Individual state variables:**

    - Household: K
    - Incumbent plant: new value of z.

- **Aggregate state variables:**

    - K, μ $(= [\mu_1, \mu_2 ..... \mu_N])$

Producer Bellman Equations in Steady State

- Incumbent plant with new productivity realization z^i:

$$V_i^p = \max\{0, \max_{h,k}[z^i f(k, h) - r^* k - w^*(h + \bar{h}_f) + \frac{1}{1+r^*-\delta} \sum_j V_j^p Q_{ij}]\}$$

The value of entry in steady state is given by:

$$-w^* \bar{h}_e + \frac{1}{1+r^*-\delta} \sum_j V_j^p G_j$$

The free entry condition reads:

$$\frac{1}{1+r-\delta} \sum_j V_j^p G_j = w^* \bar{h}_e$$

Completely standard:

$$V^h(K) = \max_{C, H, K'} \left\{ u(C, 1 - H) + \beta V^h(K') \right\}$$

$$\text{s.t. } C + K' - (1 - \delta)K = w^* H + r^* K + \Pi^*$$

$$0 \leq H \leq 1, C \geq 0, K' \geq 0$$

Steady State Equilibrium: Key Objects:

- prices: w^* and r^*

- aggregate capital and labor: K^*, H^*

- $\mu^*(z)$: invariant distribution of plants on Z

- E^*: mass of entry each period

- $x^*(z)$: exit rule (which under mild conditions will be described by a reservation productivity level z^*)

- $h^*(z)$: labor allocation to plant with shock z

- $k^*(z)$: capital allocation to plant with shock z

Steady State Equilibrium: Definition

$V^h(K), C^*, H^*, K^*, V^p(z), x^*(z), h^*(z), k^*(z), E^*, \Pi^*, \mu^*(z), w^*, r^*$ s.t.:

1. (Household optimization) V^h solves household Bellman equation, C^*, H^*, K^* are optimal choices.

2. (Firm Optimization)

   2.1 V satisfies Bellman equation for incumbents, decision rules are optimal.

   2.2 Free entry condition holds.

3. μ^*, K^* is a fixed point for law of motion.

4. Markets clear.

Solving for Steady State Equilibrium

1. Find r^*

Household side is same as in one sector growth model; standard arguments imply:

$$r^* = \frac{1}{\beta} - (1 - \delta)$$

2. Find w^*

Given r^*, solve for V^p as a function of w^*. Note: V^p is decreasing in w^*.

Use free entry condition to pin down w^*. (Unique since LHS is decreasing and RHS is increasing in w^*).

3. Find E^*

w^* implies indifference to entering. Still need to find E^*.

Stationary dist. of plants on Z is linear in E given r^*, w^*.

$\rightarrow$All aggregates are linear in entry.

Find E such that the households static FOC holds:

$$\frac{u_2(C, 1 - H)}{u_1(C, 1 - H)} = w^*$$

Note: u only affects scale of activity; TFP determined solely by steps 1,2.

1. differences in $\mu^*(z)$ due to differences in G
2. differences in $\mu^*(z)$ due to differences in Q
3. differences in $\mu^*(z)$ due to differences in $x^*(z)$
4. differences in $h^*(z)$ and $k^*(z)$ for a given $\mu^*(z)$

These channels suggest potential mechanisms for TFP differences across countries:

- Distortions that affect the ability/incentive of firms to adopt frontier technologies show up in (1) and (2).

- Distortions that prevent resources from being directed to their most productive uses show up in (3) and (4).

The Misallocation Literature

Two simple motivating observations:

1. In an economy with producers that are heterogeneous in productivity, the allocation of inputs across producers will influence (measured) aggregate TFP.

2. There are countless examples of government actions that serve to distort the allocation of inputs across producers. Moreover, it is plausible that these types of actions are more prevalent in poorer economies. (Think of corruption, crony capitalism.)

This suggests that government actions that serve to "misallocate" resources across producers could be an important determinant of aggregate TFP.

Restuccia and Rogerson (RED 2008)

- A collection of hypothetical calculations in a Hopenhayn model to explore scope of TFP effects due to misallocation.

- One approach: pick specific examples of government policies/actions, measure them and compute their effect on misallocation and aggregate TFP.

- Examples might include heterogeneous tariff/tax rates, size restrictions, subsidies to SOEs etc.....

- Issue: Largest sources of misallocation might come from hard to measure actions, e.g., corruption/crony capitalism.

- Restuccia and Rogerson use idiosyncratic producer tax/subsidy rates to broadly reflect factors that generate misallocation.

Benchmark Model

- Simple version of Hopenhayn model in SS, calibrated to match features of US economy.

- Features:

  - Inelastic labor supply
  - $f(k, h) = k^{\alpha} n^{\gamma}$
  - $\bar{h}_f = 0$; (exit via $z^0 = 0$, an absorbing state)
  - λ : probability of transitioning to z^0 from each of other z^is each period.

Calibration

Need to assign: Z, Q, G, λ, α, γ, δ, $\bar{h}_e$, β.

$\beta = 0.96 \rightarrow$ ss real interest rate of 4%

$\alpha + \gamma = 0.85 \rightarrow$ typical value, but rationale somewhat thin.

$\alpha = 0.283$, $\gamma = 0.567 \rightarrow$ capital, labor shares of 1/3 and 2/3.

$\delta = 0.08 \rightarrow$ investment share of 0.20 in ss.

z^1 normalized to 1, $h^*(z^1)$ interpreted as one employee.

$z^N = 3.98 \rightarrow h^*(z^N) = 10000 * h(z^1)$. $N = 100$ (equal log spacing).

Q is $(1-\lambda) * I$ on $\{z^1, ...z^N\}$; $\lambda = .10 \rightarrow$ job destruction rate is 10%.

G chosen to match dist. of establishment size in US.

Hypothetical Calculations

- Introduce random idiosyncratic tax rates, τ_i on value added of each producer i.

- Tax rate is only realized post-entry, may be conditioned on realized productivity, and remains constant thereafter.

- Two cases: **uncorrelated distortions** and **correlated distortions.**

  - Uncorrelated distortions: τ is independent of z.

  - Correlated: τ is correlated with z.

Calculations With Uncorrelated Distortions

- Assume $\tau_i \in \{\tau^+, \tau^-\}$ with $\tau^- < 0 < \tau^+$.

- Upon entry, half of entrants draw τ^+ and half draw τ^-.

- Given τ^+, we choose τ^- so that steady state capital is unaffected. (Recall that labor is supplied inelastically. $\rightarrow$ Steady state aggregate inputs are unchanged from benchmark.)

- They consider 4 values for τ^+: .10, .20, .30 and .40.

- They also consider scenarios where the probabilities are are something other than 50-50.

Results: Uncorrelated Distortions

Table 3
Effects of idiosyncratic distortions—uncorrelated case

Variable	τ_t			
	0.1	0.2	0.3	0.4
Relative Y	0.98	0.96	0.93	0.92
Relative TFP	0.98	0.96	0.93	0.92
Relative E	1.00	1.00	1.00	1.00
Y_s/Y	0.72	0.85	0.93	0.97
S/Y	0.05	0.08	0.09	0.10
τ_s	0.06	0.09	0.10	0.11

Table 4
Relative TFP—uncorrelated distortions

Fraction of establishments taxed (%):	τ_t			
	0.1	0.2	0.3	0.4
90	0.92	0.84	0.78	0.74
80	0.95	0.89	0.84	0.81
60	0.98	0.94	0.91	0.89
50	0.98	0.96	0.93	0.92
40	0.99	0.97	0.95	0.94
20	1.00	0.99	0.98	0.97
10	1.00	0.99	0.99	0.99

Correlated Distortions

- Repeat the previous exercise, but now assume:
  - that all producers below the median z get τ^-
  - all producers above the median z get τ^+

(i.e., tax most productive producers, subsidize least productive producers).

- As before, given τ^+, τ^- is chosen so as to leave steady state K unchanged.

Results: Correlated Distortions

Table 5
Effects of idiosyncratic distortions—correlated case

Variable	τ_t			
	0.1	0.2	0.3	0.4
Relative Y	0.90	0.80	0.73	0.69
Relative TFP	0.90	0.80	0.73	0.69
Relative E	1.00	1.00	1.00	1.00
Y_s/Y	0.42	0.67	0.83	0.92
S/Y	0.17	0.32	0.43	0.49
τ_s	0.40	0.48	0.52	0.53

Table 6
Relative TFP—correlated distortions

Fraction of establishments taxed (%):	τ_t			
	0.1	0.2	0.3	0.4
90	0.81	0.66	0.56	0.51
80	0.84	0.70	0.62	0.57
60	0.88	0.77	0.69	0.65
50	0.90	0.80	0.73	0.69
40	0.92	0.82	0.76	0.72
20	0.95	0.89	0.84	0.81
10	0.97	0.92	0.88	0.86

Discussion/Summary

- Purely random misallocation likely not that costly.

- Correlated distortions have larger impacts onTFP.

- Correlated distortions are a larger distortion. (Hopenhayn AEJ Macro 2018).

- Subsidizing most productive and taxing least productive is also TFP-reducing, but effects are much smaller because less scope for misallocation.

Measuring Misallocation and Its Effects

Two general approaches:

- Direct

- Indirect

Direct approach: the one we described previously–pick a potential source of misallocation, measure it and use a model to assess its consequences.

Indirect approach: measure the amount of misallocation without identifying the underlying source

The Indirect Approach: Basic Idea

- The indirect approach identifies misallocation from its symptoms.

- What are the key symptoms?

- Consider the static version of Hopenhayn with heterogeneous productivity, labor as the only input, $y_i = z_i h_i^\alpha$, where $0 < \alpha < 1$, and N producers.

  - Efficiency dictates equalization of marginal products:

$$\alpha z_i h_i^{\alpha-1} = \alpha z_j h_j^{\alpha-1}$$

- It follows that if we find violations of this condition then we have found misallocation.

- The bigger the violation, the bigger the misallocation.

- But how to identify violations?

- Given our functional form, the previous condition (equalization of MPL) becomes:

$$\frac{y_i}{h_i} = \frac{y_j}{h_j}$$

- Given data on y_i and h_i we can identify whether misallocation exists.

- But in fact we can do much more than just infer whether there is misallocation.

Inferring the Effect of Misallocation on TFP

- Suppose $y_i = z_i h_i^\alpha$ with α known and you have data on $\{(y_i, h_i), i = 1..N\}$.

- Define $Y = \sum_i y_i$, $H = \sum_i h_i$. Define measured TFP as Y/H.

- If we know the z_i we can solve for the level of output if labor were allocated efficiently. Denote this as Y^*.

- Can solve for z_i as: $z_i = y_i / h_i^\alpha$ given data on (y_i, h_i).

- Y/Y^* is the effect of misallocation on aggregate TFP.

- Implicitly, if y_i / h_i is not constant, $Y^* > Y$.

Conceptual Issue With This Approach

- We will discuss this in more detail after we discuss Hsieh-Klenow (2009), but the key issue is the following question.

- If you write down a model and it does not fit some aspect of the data, how do we interpret this?

- Does the above exercise detect misallocation or model misspecification?

A Practical Issue With This Approach

- Previous exercise assumed you had data on (y_i, h_i).

- Issue: Microdata typically only reports $p_i y_i$.

- Hopenhayn model assumes homogeneous output $\rightarrow p_i = p$.

- It follows that we can identify the z_i up to a scale factor.

- Scaling the z's has no impact on the efficient allocation, so can assume that $p = 1$ wlog.

- But if output is not homogeneous and prices vary, this framework cannot be used to identify the z_i.

- Need a framework that permits price heterogeneity$\rightarrow$Melitz.

Misallocation in One Factor Melitz Model

- Single household, preferences $u(C)$, labor H supplied inelastically.

- N producers, N taken as given.
    - $y_i = z_i h_i$
    - $C = Y = [\sum_{i=1}^N y_i^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}$

- Idiosyncratic taxes on revenue: τ_i

- Competitive labor market, competitive final goods producer, monopolistic competition among producers of varieties.

Identifying Misallocation in One Factor Melitz Model

- If $\tau_i = 0 \rightarrow$ equilibrium allocation is efficient.
- Producer i solves:

$$\max_{h_i} p_i z_i h_i - w h_i \text{ s.t. } p_i = B(z_i h_i)^{-1/\sigma}$$

- This implies:

$$z_i^{\frac{\sigma-1}{\sigma}} h_i^{-\frac{1}{\sigma}} = z_j^{\frac{\sigma-1}{\sigma}} h_j^{-\frac{1}{\sigma}} \text{ for all } i,j \rightarrow \frac{p_i y_i}{h_i} = \frac{p_j y_j}{h_j}$$

- i.e., efficiency $\rightarrow$ average revenue products are equated.

- Gives a simple condition for identifying misallocation.

Computing Effect of Misallocation

- To do this we need to uncover the z_i given data on $p_i y_i$ and h_i.

- Trick: Demand system lets us to do this up to a scale factor.

- Demand system implies:

$$p_i = B y_i^{-1/\sigma} \rightarrow p_i y_i = B y_i^{\frac{\sigma-1}{\sigma}} = B(z_i h_i)^{\frac{\sigma-1}{\sigma}}$$

- But now we can infer z_i using only data on $p_i y_i$ and h_i:

$$z_i = \tilde{B} \frac{(p_i y_i)^{\frac{\sigma}{\sigma-1}}}{h_i} \text{ where } \tilde{B} = B^{\frac{\sigma}{1-\sigma}}$$

Computing Effect of Misallocation cont'd

- Need to compute Y/Y^* knowing the z_i up to a scale factor.

- Key point: efficient labor allocation depends on relative z's:

$$\frac{h_i}{h_j} = \left[\frac{z_i}{z_j}\right]^{1-\sigma}, \ \sum h_i = H \rightarrow h_i^* = \frac{z_i^{1-\sigma}}{\sum z_j^{1-\sigma}}H$$

- Let h_i denote observed allocation and h_i^* denote the efficient allocation.

$$Y = \left[\sum (z_i h_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \frac{1}{\tilde{B}}\left[\sum (\tilde{B}z_i h_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

$$Y^* = \left[\sum (z_i h_i^*)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \frac{1}{\tilde{B}}\left[\sum (\tilde{B}z_i h_i^*)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

- So do not need to know $\tilde{B}$ to compute Y/Y^*.

Recovering the Distortions

- Without taxes:

$$z_i^{\frac{\sigma-1}{\sigma}} h_i^{-\frac{1}{\sigma}} = z_j^{\frac{\sigma-1}{\sigma}} h_j^{-\frac{1}{\sigma}} \rightarrow \frac{p_i y_i}{h_i} = \frac{p_j y_j}{h_j}$$

- With idiosyncratic taxes on revenue we have:

$$(1-\tau_i)z_i^{\frac{\sigma-1}{\sigma}} h_i^{-\frac{1}{\sigma}} = (1-\tau_j)z_j^{\frac{\sigma-1}{\sigma}} h_j^{-\frac{1}{\sigma}} \rightarrow \frac{(1-\tau_i)p_i y_i}{h_i} = \frac{(1-\tau_j)p_j y_j}{h_j}$$

- Which implies:

$$\frac{(1-\tau_i)}{(1-\tau_j)} = \frac{p_j y_j / h_j}{p_i y_i / h_i}$$

- Can only identify the τ_i up to a scale factor–a uniform tax on all producers does not create misallocation.

Taking as given the structure that we have imposed (including a value for σ), and given data on $p_i y_i$ and h_i, we have shown:

1. how to infer the TFP loss due to misallocation

2. how to uncover the implicit distortions causing misallocation.

Generalization to Multiple Factors

- Generalize previous analysis:

$$y_i = z_i k_i^\alpha h_i^{1-\alpha}$$

- Without idiosyncratic distortions, equilibrium is efficient and both $p_i y_i / k_i$ and $p_i y_i / h_i$ are equated across producers.

- Multiple factors $\rightarrow$ multiple margins of misallocation.

- Can consider idiosyncratic taxes on revenue, capital and labor, but only 2 can be determined.

- Useful representation is a distortion to output and a distortion to k/h.

- Can uncover z_i as before given data on $p_i y_i$, k_i and h_i.

TFPQ and TFPR

- Previous discussion highlights an issue in computing TFP in micro data.

- This has led to researchers distinguishing two different concepts of TFP.

- TFPR is a revenue based measure of TFP.

- TFPQ is a quantity based measure of TFP.

- Important point is that we fundamentally care about TFPQ, but in micro data we typically do not have good measures of quantity and so have to make do with revenue.

- This paper did two things:

  - developed the preceding methodology

  - applied it to the manufacturing sectors of China, India and the US to infer the importance of misallocation in these three contexts.

- Key finding: misallocation is quantitatively important, and much more so in China and India than in the US.
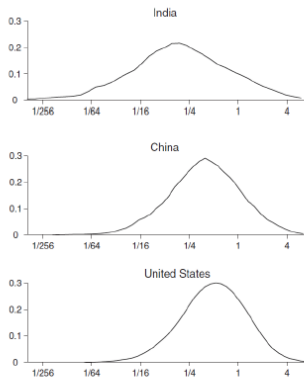
FIGURE I
Distribution of TFPQ

Properties of TFPQ, TFPR

TABLE I
DISPERSION OF TFPQ

China	1998	2001	2005
S.D.	1.06	0.99	0.95
75 − 25	1.41	1.34	1.28
90 − 10	2.72	2.54	2.44
N	95,980	108,702	211,304
India	1987	1991	1994
S.D.	1.16	1.17	1.23
75 − 25	1.55	1.53	1.60
90 − 10	2.97	3.01	3.11
N	31,602	37,520	41,006
United States	1977	1987	1997
S.D.	0.85	0.79	0.84
75 − 25	1.22	1.09	1.17
90 − 10	2.22	2.05	2.18
N	164,971	173,651	194,669

Notes. For plant i in industry s, TFPQ$_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s}(w_s L_{si})^{1-\alpha_s}}$. Statistics are for deviations of log TFPQ from industry means. S.D. = standard deviation, 74 − 26 is the difference between the 75th and 25th percentiles, and 90 − 10 the 90th vs. 10th percentiles. Industries are weighted by their value-added shares. N = the number of plants.

TABLE II
DISPERSION OF TFPR

China	1998	2001	2005
S.D.	0.74	0.68	0.63
75 − 25	0.97	0.88	0.82
90 − 10	1.87	1.71	1.59
India	1987	1991	1994
S.D.	0.69	0.67	0.67
75 − 25	0.79	0.81	0.81
90 − 10	1.73	1.64	1.60
United States	1977	1987	1997
S.D.	0.45	0.41	0.49
75 − 25	0.46	0.41	0.53
90 − 10	1.04	1.01	1.19
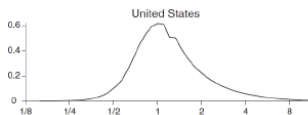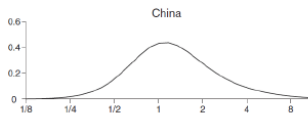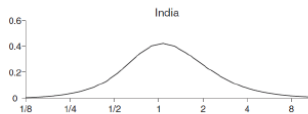
FIGURE II
Distribution of TFPR

TABLE IV

TFP GAINS FROM EQUALIZING TFPR WITHIN INDUSTRIES

China	1998	2001	2005
%	115.1	95.8	86.6
India	1987	1991	1994
%	100.4	102.1	127.5
United States	1977	1987	1997
%	36.1	30.7	42.9

Notes. Entries are $100(Y_{\text{efficient}}/Y - 1)$ where $Y/Y_{\text{efficient}} = \prod_{s=1}^{S} [\sum_{i=1}^{M_s} (\frac{A_{si}}{\bar{A}_s} \frac{\overline{\text{TFPR}}_s}{\text{TFPR}_{si}})^{\sigma-1}]^{\theta_s/(\sigma-1)}$ and $\text{TFPR}_{si} = \frac{P_{si} Y_{si}}{K_{si}^{\alpha_s} (w_{si} L_{si})^{1-\alpha_s}}$.

- Hsieh-Klenow methodology is basically an example of the "wedge" methodology that has been used elsewhere, mostly in the business cycle literature.

- Issue: Are the wedges evidence of model misspecification or primitive distortions?

  - Example 1: If k/h varies across producers, is this misallocation or evidence of heterogeneous technologies (i.e., different α's)?

  - Example 2: Distortion or measurement error?

- More generally, easy to write down a model where efficient allocations imply dispersion in average revenue products.