Will the average worker be left behind?

Xiaoyue Zhang [1] Junjie Xia [2]

[1]Tilburg University

[2]CUFE & Peking University

October 29, 2022

Motivation

- Different allocation of input is important for explaining the disparity in income per capita across countries.

- Developing countries may become richer if input can be allocated more efficiently.

Motivation: research questions

- Who benefit from this process and how?
- Will the average worker be left behind?

Motivation

Hsieh and Klenow (2009)

- Distortions cause too much or too few inputs usage.
- Nested CES and each industry is a nest.
- All the nests have the **same** demand elasticities.

Changes in labor income share from removing distortions are affected by:

- aggregation structure
- distribution of technology
- interaction with the asymmetric distortions.

Motivation

Hsieh and Klenow (2009)

- Distortions cause too much or too few inputs usage.
- Nested CES and industry $=$ nest.
- The **same** demand elasticities.

Changes in labor income from removing distortions are affected by:

- aggregation structure
- distribution of technology
- interaction with the asymmetric distortions.

But

- Difficult to define a market.
- An industry $\neq$ a nest (market).

And in developing countries

- Weaker anti-competitive regulation.
- Large dispersion in markups.

Motivation

Hsieh and Klenow (2009)

- Distortions cause too much or too few inputs usage.
- Nested CES and industry $=$ nest.
- The **same** demand elasticities.

Changes in labor income from removing distortions are affected by:

- aggregation structure
- distribution of technology
- **distribution of markups**
- interaction with the asymmetric distortions.

But

- Difficult to define a market.
- An industry $\neq$ a nest (market).

And in developing countries

- Weaker anti-competitive regulation.
- Large dispersion in markups.

Markups distribution is also important for labor share gains.

Motivation

Hsieh and Klenow (2009)

- Distortions cause too much or too few inputs usage.
- Nested CES and industry = nest.
- The **same** demand elasticities.

Changes in labor income from removing distortions affected by:

- aggregation structure
- distribution of technology
- **distribution of markups**
- interaction with the asymmetric distortions.

But

- Difficult to define a market.
- An industry $\neq$ a nest (market).

And in developing countries

- Weaker anti-competitive regulation.
- Large dispersion in markups.

Markups distribution is also important for labor share gains.

How to estimate gains with noisy market info and different demand elasticities?

Key ideas

- The underlying market structure is a system of latent clusters (nests) of firms.
- Infer the latent nests using observed firm characteristics.
- Estimate industry-specific production elasticities, firm-specific distortions, nest-specific demand elasticities.

Key ideas

- The underlying market structure is a system of latent clusters (nests) of firms.
- Infer the latent nests using observed firm characteristics.
  - Industry category + revenue-cost ratio (today)
  - More variables: location, number of patents, age, ownership, etc. (future work)
- Estimate industry-specific production elasticities, firm-specific distortions, nest-specific demand elasticities.

Main results

Applying this method to 2005 Chinese firm-level data, we find:

- 90% of the industries are better modelled as having more than one nest/market;
- K is more often subsidized than L.
- Removing input distortions, reallocate:
  - L to markets with lower markups: larger gains to L,
  - K to markets with higher markups: smaller gains to K,
  - K and L to bigger markets with higher demand for K and L: higher gains.
- The average worker benefits disproportionally more.

Merits

- The pattern between markups and market shares within each industry can be arbitrary. (Atkeson and Burstein (2008), Haltiwanger et al. (2018), Peters (2020), Ruzic and Ho (2021), Liang (2021), and Gupta (2021)) proxy

- Allow an industry to have latent markets.

- Easy to implement

Misallocation: Restuccia and Rogerson (2008) and Hsieh and Klenow (2009).

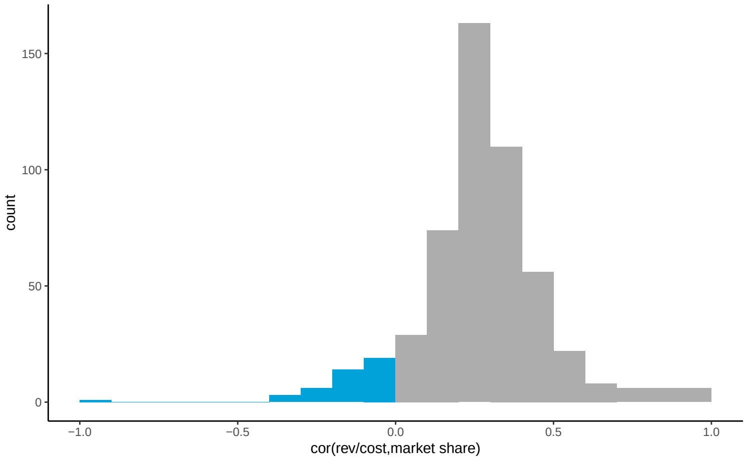
Labor share and markups: Autor et al. (2020).

Sources of TFPR variation: Haltiwanger et al. (2018), David and Venkateswaran (2019), and Bils et al. (2020).

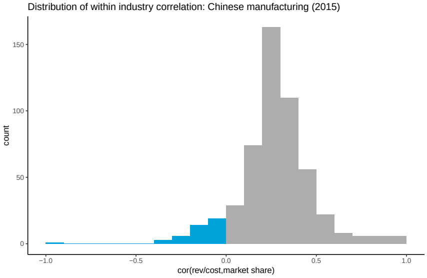
Latent cluster structure: Bonhomme and Manresa (2015), Bonhomme et al. (2022).

Industry $\neq$ market

Distribution of within industry correlation: Chinese manufacturing (2015)

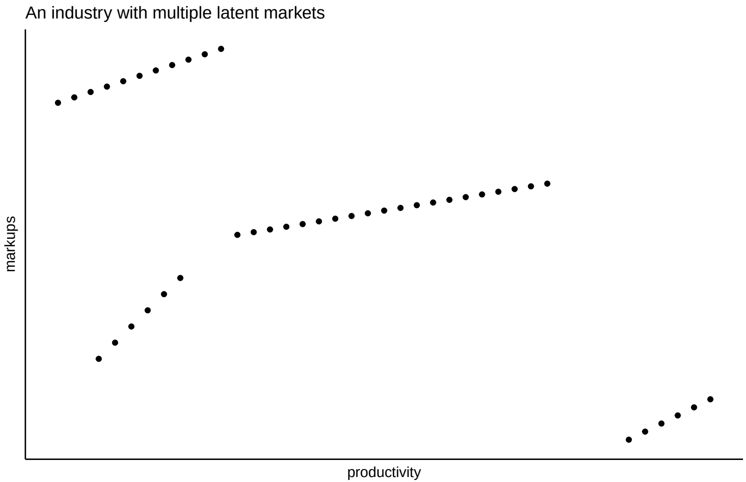


Industry $\neq$ market

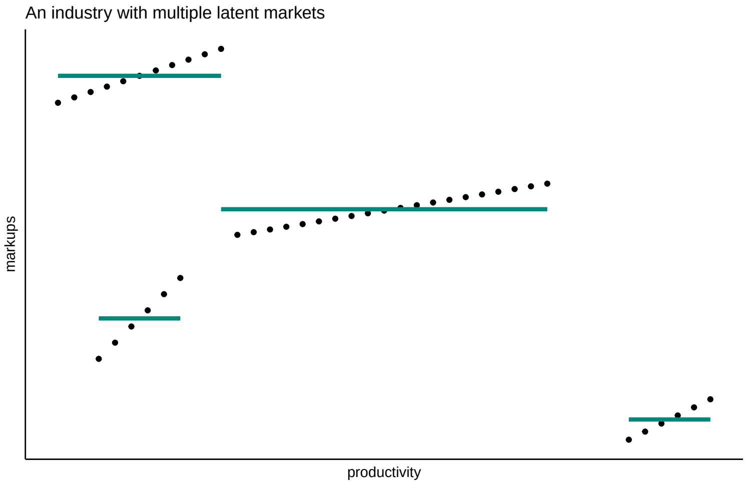


- Various levels of pass-through:
  - complete pass-through, incomplete pass-through, etc.
- An industry $\neq$ a market:
  - multiple markets exist in an industry (today),
  - firms from different industries belong to the same market (future work).

Latent market structure: an illustrative example



Latent market structure: an illustrative example



Model

Demand: nested CES that allows industries to have multiple nests with different demand elasticities.

Supply: Cobb-Douglas production $Y_i = A_i K_i^{\alpha_s^K} L_i^{\alpha_s^L}$ with

- industry-specific production elasticities,
- firm-specific positive or negative distortions (Restuccia and Rogerson (2008)),
- firm-specific productivity.

$$\Pi_i = P_i Y_i - (R(1+\tau_i^K)K_i + w(1+\tau_i^L)L_i)$$

Model

Demand: nested CES that allows industries to have multiple nests with different demand elasticities.

Supply: Cobb-Douglas production $Y_i = A_i K_i^{\alpha_s^K} L_i^{\alpha_s^L}$ with

- industry-specific production elasticities,
- firm-specific positive or negative distortions (Restuccia and Rogerson (2008)),
- firm-specific productivity.

$$\Pi_i = P_i Y_i - (R(1+\tau_i^K)K_i + w(1+\tau_i^L)L_i)$$

Firms' FOC:

$$\log(1+\tau_i^L) = -log(\frac{\epsilon_g}{\epsilon_g - 1}) + \log(\alpha_s^L) - \log\left(\frac{wL_i}{P_i Y_i}\right)$$

Aggregate input shares with τ

$$\frac{wL}{PY} = \sum_g \frac{P_g Y_g}{PY} \sum_{i \in g} \frac{wL_i}{P_i Y_i} \frac{P_i Y_i}{P_g Y_g}$$

$$= \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{distribution across nests}} \cdot \underbrace{\sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g}\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{distribution within nests}}$$

Changes in input shares

$$\frac{w^* L}{P^* Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g} \left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{reallocation within nests}} \right]$$

- Due to within-nest reallocation: the level of τ_i^L and the joint distribution of $A_i$and τ_i^L within nests.
- The change in nest-level labor share

Changes in input shares

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g} \left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{reallocation within nests}} \right]$$

- Due to across-nest reallocation: the joint distribution of nest-level changes and the aggregation structure:
  - technology α_g^L,
  - demand elasticities ϵ_g,
  - the importance of the nests β_g.

Changes in input shares

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g} \left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{reallocation within nests}} \right]$$

- Due to across-nest reallocation: the joint distribution of nest-level changes and the aggregation structure:
- Nest g's labor share change is more important if β_g, α_g and $\frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}$ are higher.

Changes in input shares

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g} \left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{reallocation within nests}} \right]$$

- Decompose the change into nest-level change and the aggregation structure.
- The part due to nest-level changes:
  - no variation in α_g^L, ϵ_g, and β_g
  - holding the nest-level change constant.

Changes in input shares

$$\frac{w^* L}{P^* Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1 + \bar{\tau}_g^L}}_{\text{reallocation within nests}} \right]$$

- Decompose the change into nest-level change and the aggregation structure.
- The part due to nest-level changes:
  - no variation in α_g^L, ϵ_g, and β_g
  - holding the nest-level change constant.

Changes in input shares

$$\frac{w^* L}{P^* Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta \alpha^L \frac{\epsilon - 1}{\epsilon \mathbb{E}[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1 + \bar{\tau}_g^L}}_{\text{nest-level change fixed}} \right]$$

- Decompose the change into nest-level change and the aggregation structure. xxxx
- The part due to nest-level changes:
  - no variation in α_g^L, ϵ_g, and β_g
  - holding the nest-level change constant.

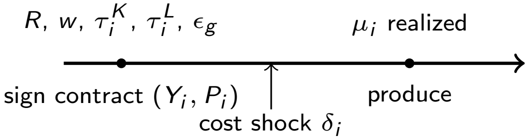
$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \sum_{i \in g} \frac{1}{1+\tau_i^L} \frac{\left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}{\sum_{i \in g} \left(\frac{A_i}{(1+\tau_i^L)^{\alpha_g^L}}\right)^{\frac{\epsilon_g - 1}{\epsilon_g - (\epsilon_g - 1)\alpha_g^L}}}}_{\text{reallocation within nests}} \right]$$

Challenges:

- Better not to use US as a benchmark:
  - differences in technology between US and China will be treated as input distortions,
  - fine when talking about TFP gains,
  - but create a systematic bias for labor share change.
- Fit a nested CES demand to the firm-level revenue-cost ratios in the data.

Estimation

- A parametric assumption about the input distortions: estimate α_s^K, α_s^L
  - Assumption: most firms have 0 distortions.
- Idiosyncratic cost shock and price rigidity to explain the remaining variation in revenue-cost ratio



Estimation

- A parametric assumption about the input distortions: estimate α_s^K, α_s^L
  - Assumption: most firms have 0 distortions.
- Idiosyncratic cost shock and price rigidity to explain the remaining variation in revenue-cost ratio

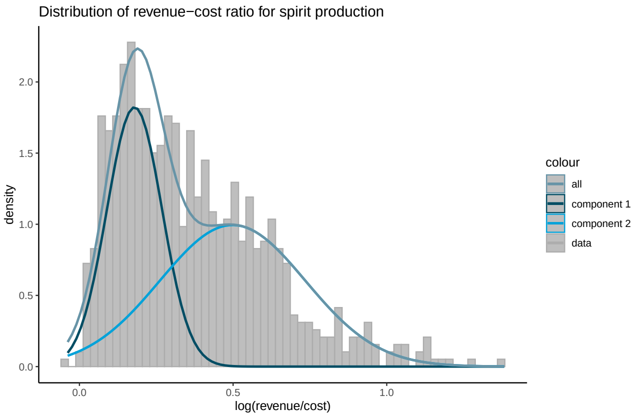
Realized profits:

$$\Pi_i = P_i Y_i - (R(1+\tau_i^K)K_i + w(1+\tau_i^L)L_i)e^{\delta_i}$$

Realized markups (revenue-cost ratio):

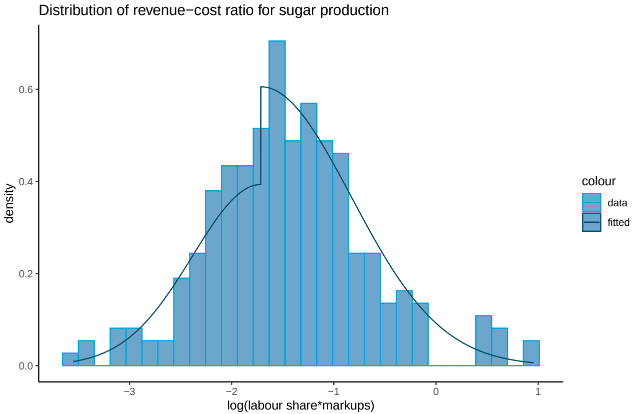
$$\mu_i = \frac{\epsilon_g}{\epsilon_g - 1}\frac{\mathbb{E}[e^{\delta_i}]}{e^{\delta_{ig}}} - 1$$

Markups distribution



formula

Labor share distribution



$$\log(\text{labor share}_i * \mathbb{E}[\text{markups}_i]) = \log(\alpha_s^L) - \log(1 + \tau_i^L)$$

Likelihood function Estimation

Data

Chinese Annual Firm-Level Survey Data (2005) from NBS.

Statistic	N	Mean	Min	Pctl(25)	Median	Pctl(75)	Max
value added	229,241	13,814.46	122	2,517	5,377	13,250	277,908
K	229,241	16,366.41	83.76	1,620.23	4,211.66	12,151.88	515,954.20
wL	229,241	2,730.73	80	583	1,188	2,665	78,956
revenue	229,241	50,184.74	2	9,500	19,457	45,994	11,041,153
cost	229,241	43,075.61	1	7,935	16,481	39,072	10,757,115
profits	229,241	2,370.47	−292,087	72	480	1,815	415,879
revenue/cost	229,241	1.21	0.81	1.08	1.14	1.25	4.68
wL/value added	229,241	0.32	0.01	0.12	0.23	0.42	3.15

Results: estimated parameters

two types	N	Mean	Min	Pctl(25)	Median	Pctl(75)	Max
No	61	23	2	6	15	27	237
Yes	462	494	12	118	256	544.500	9,947

	N	Mean	St. Dev.	Pctl(10)	Pctl(25)	Median	Pctl(75)	Pctl(90)
$\mathbb{E}_g[\mu_i + 1]$	985	1.30	0.25	1.11	1.14	1.22	1.39	1.57
σ_g	985	6.33	3.64	2.77	3.59	5.45	8.32	10.48
$\mathbb{E}_g[e_i^\delta]$	985	1.01	0.02	1	1	1.01	1.02	1.03
ex-ante $P_g[\bar{s}]$	928	0.66	0.22	0.27	0.59	0.73	0.82	0.88
α_K	523	0.16	0.17	0.04	0.06	0.09	0.19	0.36
α_L	523	0.39	0.23	0.13	0.21	0.33	0.57	0.76
scale	523	0.55	0.31	0.22	0.32	0.48	0.75	0.95

Dispersion of markups within an industry

Results: Labor and capital income share (%)

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1+\bar{\tau}_g^L}}_{\text{nest-level change fixed}} \right]$$

(a) Benchmark

	observed	predicted	change
L	19.76	27.2	7.44
K	11.86	10.77	-1.09
L+K	31.62	37.97	6.35

(b) Nest-level τ fixed & homo α, β, and σ

	observed	predicted	change
L	24.69	26.92	2.23
K	44.57	26.92	-17.65
L+K	69.26	53.85	-15.42

Results: Labor and capital income share (%)

$$\frac{w^* L}{P^* Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta\alpha^L \frac{\epsilon - 1}{\epsilon \mathbb{E}[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1 + \bar{\tau}_g^L}}_{\text{nest-level change fixed}} \right]$$

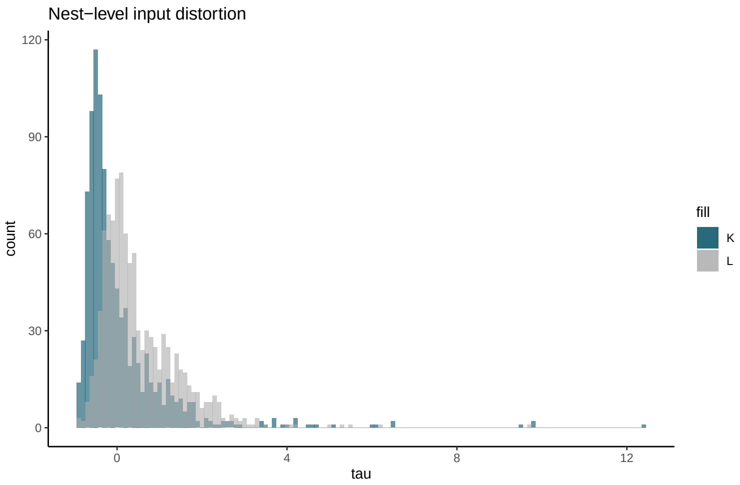
(a) Benchmark

	observed	predicted	change
L	19.76	27.2	7.44
K	11.86	10.77	-1.09
L+K	31.62	37.97	6.35

(b) Nest-level τ fixed & homo α, β, and σ

	observed	predicted	change
L	24.69	26.92	2.23
K	44.57	26.92	-17.65
L+K	69.26	53.85	-15.42

Explanation: homogeneous α, β, and σ



- Labor income share increases but capital income share decreases.

Explanation: compared to benchmark

Table: Correlations behind the cross-nest reallocation

$cor(\bar{\tau}^L_g, \beta_g \alpha^L_g \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]})$	$cor(\bar{\tau}^K_g, \beta_g \alpha^L_g \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]})$
0.05	0.04

- The labor share and capital share gains are larger in the benchmark

Results: Labor and capital income share (%)

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha_g^L \frac{\epsilon - 1}{\epsilon \mathbb{E}[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1 + \bar{\tau}_g}}_{\text{nest-level change fixed}} \right]$$

(a) Benchmark

	observed	predicted	change
L	19.76	27.2	7.44
K	11.86	10.77	-1.09
L+K	31.62	37.97	6.35

(b) Nest-level τ fixed & homo ϵ

	observed	predicted	change
L	18.54	25.43	6.88
K	11.29	10.26	-1.03
L+K	29.83	35.69	5.85

(a) Benchmark

	observed	predicted	change
L	19.76	27.2	7.44
K	11.86	10.77	-1.09
L+K	31.62	37.97	6.35

(b) Nest-level τ fixed & homo ϵ

	observed	predicted	change
L	18.54	25.43	6.88
K	11.29	10.26	-1.03
L+K	29.83	35.69	5.85

(a) Nest-level τ fixed & homo α, β, and σ

	observed	predicted	change
L	24.69	26.92	2.23
K	44.57	26.92	-17.65
L+K	69.26	53.85	-15.42

Explanation

Table: Correlations behind the cross-nest reallocation: homo α and β

$cor(\bar{\tau}_g^L, \beta_g \alpha_g^L)$	$cor(\bar{\tau}_g^K, \beta_g \alpha_g^L)$
0.05	0.04

- Variations in α and β raise the gains for labor and capital shares

$cor(\epsilon_g/(\epsilon_g - 1), L^*/L)$	$cor(\epsilon_g/(\epsilon_g - 1), K^*/K)$
-0.09	0.02

- Labour is more likely reallocated to low-markup nests.
- Capital is more likely reallocated to high-markup nests.

Results: Labor and capital income share (%)

$$\frac{w^*L}{P^*Y^*} - \frac{wL}{PY} = \sum_g \underbrace{\beta_g \alpha^L \frac{\epsilon_g - 1}{\epsilon_g \mathbb{E}_g[e^{\delta_i}]}}_{\text{reallocation across nests}} \cdot \left[\underbrace{1 - \frac{1}{1+\bar{\tau}_g}}_{\text{nest-level change fixed}} \right]$$

(a) Benchmark

	observed	predicted	change
L	19.76	27.2	7.44
K	11.86	10.77	-1.09
L+K	31.62	37.97	6.35

(b) Nest-level τ fixed & homo α

	observed	predicted	change
L	25.55	29.20	3.65
K	42.63	29.20	-13.43
L+K	68.18	58.40	-9.78

Explanation

Table: Correlations behind the cross-nest reallocation: homo α

$cor(\alpha_g^L, L^*/L)$	$cor(\alpha_g^K, K^*/K)$
0.68	0.62

- K and L are reallocated to nests with higher α.

Conclusion

- Build on HK's model to study changes in input shares when removing input distortions.
- Decompose the changes into two parts: caused by within-nest and by across-nest reallocation.

Apply the model to Chinese manufacturing firms (2005) and find:

- 90% of the industries are better modelled as having more than one nest/market.
- in general, firms use too little L and too much K:
  - L experience higher idiosyncratic usage cost while K is often subsidized
- when removing the input distortions, the across-nest reallocation raises the gains to L and K:
  - K and L are reallocated to larger markets with higher demand for them.
  - L is reallocated to lower-markups markets but K to higher-markups ones.
- The average worker benefits disproportionally more.

References

Atkeson, A. and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *The American Economic Review*, 98, 1998–2031, publisher: American Economic Association.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 135, 645–709.

Baqaee, D. R. and E. Farhi (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Bils, M., P. J. Klenow, and C. Ruane (2020): "Misallocation or Mismeasurement?" Working Paper 26711, National Bureau of Economic Research.

Bonhomme, S., T. Lamadon, and E. Manresa (2022): "Discretizing Unobserved Heterogeneity," *Econometrica*, 90, 625–643.

Bonhomme, S. and E. Manresa (2015): "GROUPED PATTERNS OF HETEROGENEITY IN PANEL DATA," *Econometrica*, 83, 1147–1184.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2014): "Challenges of Working with the Chinese NBS Firm-Level Data," *China Economic Review*, 30, 339–352.

David, J. M. and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109, 2531–2567.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 135, 561–644.

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

Edmond, C., V. Midrigan, and D. Y. Xu (2019): "How Costly Are Markups," *working paper*.

Feenstra, R. C. and D. E. Weinstein (2017): "Globalization, Markups, and US Welfare," *Journal of Political Economy*, 125, 1040–1074, publisher: The University of Chicago Press.

Gupta, A. (2021): "Demand for Quality, Variable Markups and Misallocation: Evidence from India," *working paper*.

Haltiwanger, J., R. Kulick, and C. Syverson (2018): "Misallocation Measures: The Distortion That Ate the Residual," Working Paper 24199, National Bureau of Economic Research.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 46.

Klette, T. J. and Z. Griliches (1996): "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous," *Journal of Applied Econometrics*, 11, 343–361.

Liang, Y. (2021): "Misallocations and Markups: Evidence from Indian Manufacturing," *working paper*, 59.

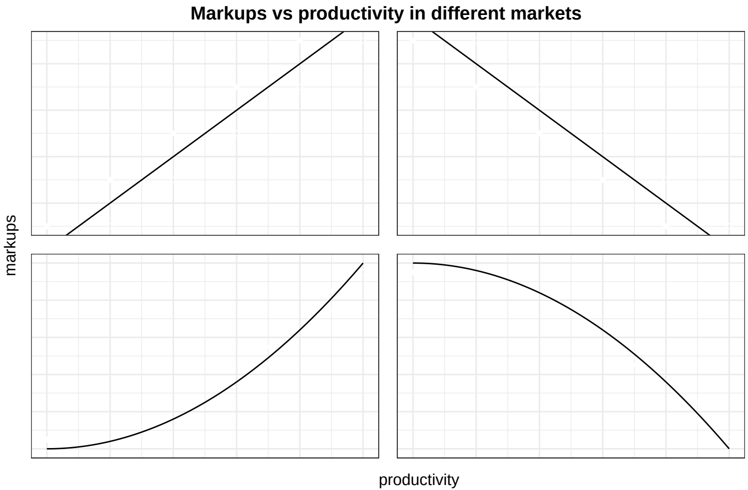
Peters, M. (2020): "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 88, 2037–2073.

Restuccia, D. and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707–720.

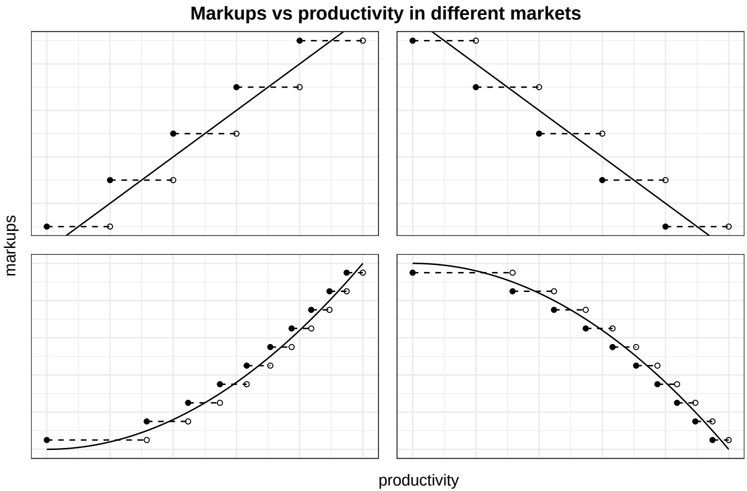
Ruzic, D. and S.-J. Ho (2021): "Returns to Scale, Productivity Measurement, and Trends in U.S. Manufacturing Misallocation," *The Review of Economics and Statistics*, 1–47.

Thank you!

Different relations between markups and productivity



Different relations between markups and productivity



return

Distributions of exogenous variables

Cost shock are i.i.d:

$$\delta_{ig} \sim \mathcal{N}(0, \sigma_g)$$

τ_i^K and τ_i^L are i.i.d within industry and independent across industries:

$$\log(\tau_i^K + 1) \sim \begin{cases} 2\kappa^K \mathcal{N}(0, \sigma_-^K) \text{ , if } \tau_i^K < 0 \\ (2 - 2\kappa^K)\mathcal{N}(0, \sigma_+^K) \text{ , if } \tau_i^K > 0 \end{cases}$$

$$\log(\tau_i^L + 1) \sim \begin{cases} 2\kappa^L \mathcal{N}(0, \sigma_-^L) \text{ , if } \tau_i^L < 0 \\ (2 - 2\kappa^L)\mathcal{N}(0, \sigma_+^L) \text{ , if } \tau_i^L > 0 \end{cases}$$

$\tau_i^f > 0$ for $f \in K, L$: firms hire capital or labor at a price higher than market price.
$\tau_i^f < 0$ for $f \in K, L$: firms hire at a price lower than market price.
κ^K, κ^L, σ_+^K, σ_-^K, σ_+^L, and σ_-^L are distribution parameters. Return

Step 2: groups and demand elasticities

Realized markups:

$$\mu_{ig} = \frac{\epsilon_g}{\epsilon_g - 1} \frac{\mathbb{E}[e^{\delta_i}]}{e^{\delta_{ig}}} - 1$$

Distribution of markups observed in an industry is a mixture of two distributions, one from the group with higher demand elasticities, $\epsilon_{\bar{s}}$, and one from the group with lower demand elasticities, $\epsilon_{\underline{s}}$.

$$\log(\mu_i + 1) \sim w_s \cdot \underbrace{\mathcal{N}\left(\log \frac{\epsilon_{\underline{s}} e^{\sigma^2_{\epsilon_{\underline{s}}}/2}}{\epsilon_{\underline{s}} - 1}, \sigma_{\epsilon_{\underline{s}}}\right)}_{\text{low-demand-elasticities group}} + (1 - w_s) \cdot \underbrace{\mathcal{N}\left(\log \frac{\epsilon_{\bar{s}} e^{\sigma^2_{\epsilon_{\bar{s}}}/2}}{\epsilon_{\bar{s}} - 1}, \sigma_{\epsilon_{\bar{s}}}\right)}_{\text{high-demand-elasticities group}}$$

First test whether there exists more than one distribution. Then use the EM algorithm to estimate w_s, $\epsilon_{\bar{s}}$, $\epsilon_{\underline{s}}$, $\sigma_{\epsilon_{\bar{s}}}$, and $\sigma_{\epsilon_{\underline{s}}}$.

Return

Likelihood function:

$$\ell(P_iY_i, K_i, L_i|\epsilon, \alpha, \sigma, \kappa, R, w) = \ell(P_iY_i, K_i, |\epsilon, \alpha, \sigma, \kappa, R, w) + \ell(P_iY_i, L_i|\epsilon, \alpha, \sigma, \kappa, R, w)$$

$$\ell(P_iY_i, K_i|\epsilon, \alpha, \sigma, \kappa, R, w) \propto \left[\log(2\kappa^K) - \log\sigma_+^K - \frac{1}{2}\left(\frac{\log(\frac{\alpha_s^K P_iY_i(\epsilon-1)/\epsilon}{RK_i\mathbb{E}[e^{\delta_i}]})}{\sigma_+^K}\right)^2\right]\mathbb{1}\left[\frac{\alpha_s^K P_iY_i(\epsilon-1)/\epsilon}{RK_i\mathbb{E}[e^{\delta_i}]} > 1\right]$$
$$+\left[\log(2-2\kappa^K) - \log\sigma_-^K - \frac{1}{2}\left(\frac{\log(\frac{\alpha_s^K P_iY_i(\epsilon-1)/\epsilon}{RK_i\mathbb{E}[e^{\delta_i}]})}{\sigma_-^K}\right)^2\right]\mathbb{1}\left[\frac{\alpha_s^K P_iY_i(\epsilon-1)/\epsilon}{RK_i\mathbb{E}[e^{\delta_i}]} < 1\right]$$

$$\ell(P_iY_i, L_i|\epsilon, \alpha, \sigma, \kappa, R, w) \propto \left[\log(2\kappa^L) - \log\sigma_+^L - \frac{1}{2}\left(\frac{\log(\frac{\alpha_s^L P_iY_i(\epsilon-1)/\epsilon}{wL_i\mathbb{E}[e^{\delta_i}]})}{\sigma_+^L}\right)^2\right]\mathbb{1}\left[\frac{\alpha_s^L P_iY_i(\epsilon-1)/\epsilon}{wL_i\mathbb{E}[e^{\delta_i}]} > 1\right]$$
$$+\left[\log(2-2\kappa^L) - \log\sigma_-^L - \frac{1}{2}\left(\frac{\log(\frac{\alpha_s^L P_iY_i(\epsilon-1)/\epsilon}{wL_i\mathbb{E}[e^{\delta_i}]})}{\sigma_-^L}\right)^2\right]\mathbb{1}\left[\frac{\alpha_s^L P_iY_i(\epsilon-1)/\epsilon}{wL_i\mathbb{E}[e^{\delta_i}]} < 1\right]$$

Return

Extension: estimate α_s^K, α_s^L

1. Draw a grid on the domain of α_s^K and α_s^L, $(0,1) \times (0,1)$. The density of this grid affects the accuracy of estimated α_s^K and α_s^L.
2. For each point on the grid, set α_s^K and α_s^L equal to the value of this point, i.e. a guess of α_s^K and α_s^L.
3. For each industry, estimate $\widehat{\sigma_+^K}$,$\widehat{\sigma_-^K}$,$\widehat{\sigma_+^L}$,$\widehat{\sigma_-^L}$,$\widehat{\kappa^K}$,$\widehat{\kappa^L}$ according to the equations above.
4. Calculate log-likelihood for each industry at the guessed α_s^K and α_s^L.
5. Find the α_s^K and α_s^L that give the highest log-likelihood, which is the estimated capital intensity of this industry.

Return