Abstract

China's Economic Slowdown: A Firm-Level Perspective

Loren Brandt, Ruochen Dai, Gueorgui Kambourov,
Zhenjie Qian, Kjetil Storesletten and Luhang Wang

May 2026

This paper develops a new representative firm-level dataset for Chinese manufacturing from 1998 to 2019 by combining tax, registry, and inspection records and correcting for sampling biases in existing administrative data. Using a CES production framework with firm-specific distortions, the paper distinguishes labor-augmenting from Hicks-neutral productivity growth and introduces a novel decomposition of aggregate TFP growth into efficient growth and changes in misallocation. The analysis documents a sharp slowdown in Chinese manufacturing productivity growth after 2010, driven both by weaker incumbent-firm productivity growth and by a diminished contribution from firm turnover and reallocation. Before 2010, rapid productivity growth reflected declining misallocation over firms' life cycles, especially among young firms. After 2010, however, distortions affecting new firms dissipated more slowly, leading to persistent misallocation and substantially lower aggregate TFP growth.

China's Economic Slowdown: A Firm-Level Perspective

Loren Brandt[1] Ruochen Dai[2] Gueorgui Kambourov[1]
Zhenjie Qian[3] Kjetil Storesletten[4] Luhang Wang[5]

[1]U of Toronto [2]Central U of Finance and Economics [3]U of International Business and Economics [4]U of Minnesota [5]Xiamen U

EIEF

May 25, 2026

Motivation

Claim: China's growth in GDP & TFP has fallen since 2007

- Is it true?
- If so, why?
  Alternative narratives:
  1. Low-hanging fruits exhausted, slowdown in US/OECD
  2. Government policies less GDP-growth oriented
     (Anti-corruption campaign, fiscal expansion, etc.)
     - Return of firm-level distortions & barriers
       $\Rightarrow$ more misallocation, less churn & entrepreneurial activity
     - Misguided strategic priorities (industrial policy)
       size-based policies, real estate boom/bust

This Paper

- Build representative data set for Chinese manufacturing until 2019
- Study productivity growth and factor-saving bias
- Use firm-level data for bottom-up view;
  1. Identify labor-augmenting vs. Hicks-neutral growth
  2. Study role of size and churn (reallocation, entry, exit)
  3. Novel growth decomposition: efficiency vs. misallocation

Challenge: Data Availability

- National accounts are probably exaggerated
- Need representative firm-level data to calculate TFP
- Alternative firm-level data sets for China manufacturing:
  (1) **NBS**; (2) **Tax data (NTB)**; (3) **Inspection Data**
  1. NBS data have problems after 2007 and ends in 2013
  2. National Tax Bureau (NTB) data has good measurements
     ... but not representative (oversamples older, larger firms)
  3. Inspection data covers the universe of firms
     ... but few variables
- Solution: use Inspection Data to adjust sampling weights for NTB. Obtain representative firm-level data up to 2019

Description of Data Sources (2007-2019)

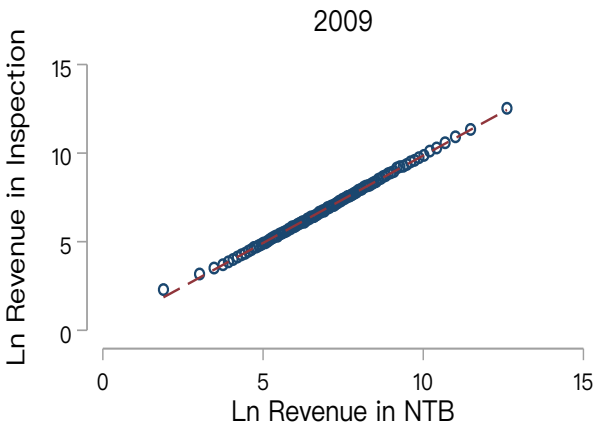
- NTB Data 2003-2019
  - : Collected by tax authorities. Repeated cross section
  - : Purpose: taxation of value-added and profits
- Registry data
  - : Universal coverage of all registered corporate entities.
  - : *Key Variables:* Entry and exit dates, location, and ownership type.
- Inspection data 2007-2019
  - : Panel data collected via mandatory annual filings to SAIC.
  - : *Key Variables:* assets, revenue, operational status

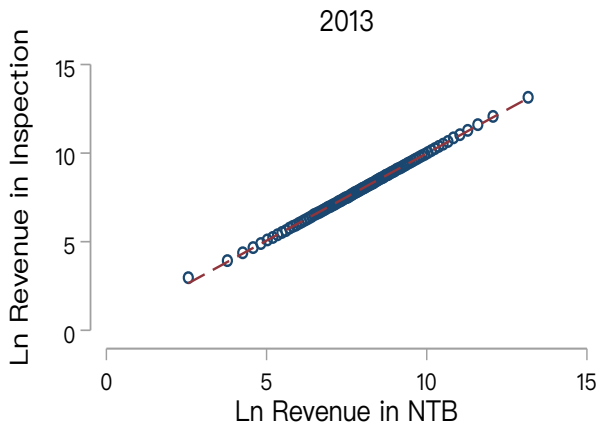
Description of Data Sources (1998-2007)

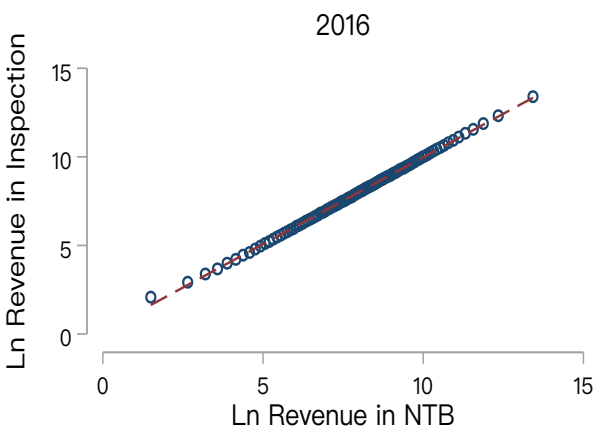
- NBS data 1998-2007
  - : Collected by National Bureau of Statistics
  - : Sample: all SOEs and all firms with revenue “above scale”

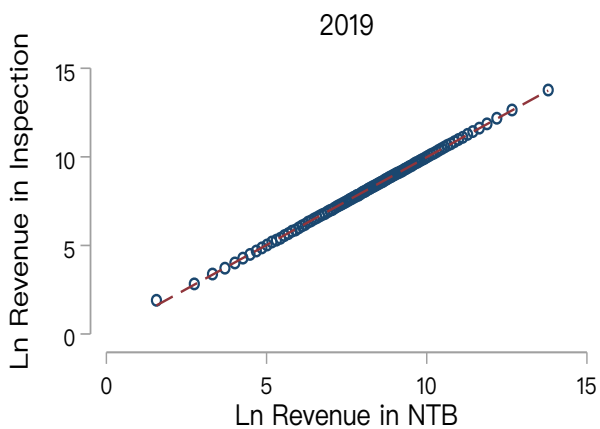
Claim: NTB & Inspection Data Have High Quality

- Match firms in NTB and Inspection Data on Legal Repr. ID, firm name, and address (match $>$ 95% of NTB firms)
- Estimate classical m.e. in revenue and assets:
  var(m.e.) $<$ 0.7% of variance in NTB, $<$ 5% for Inspection Data

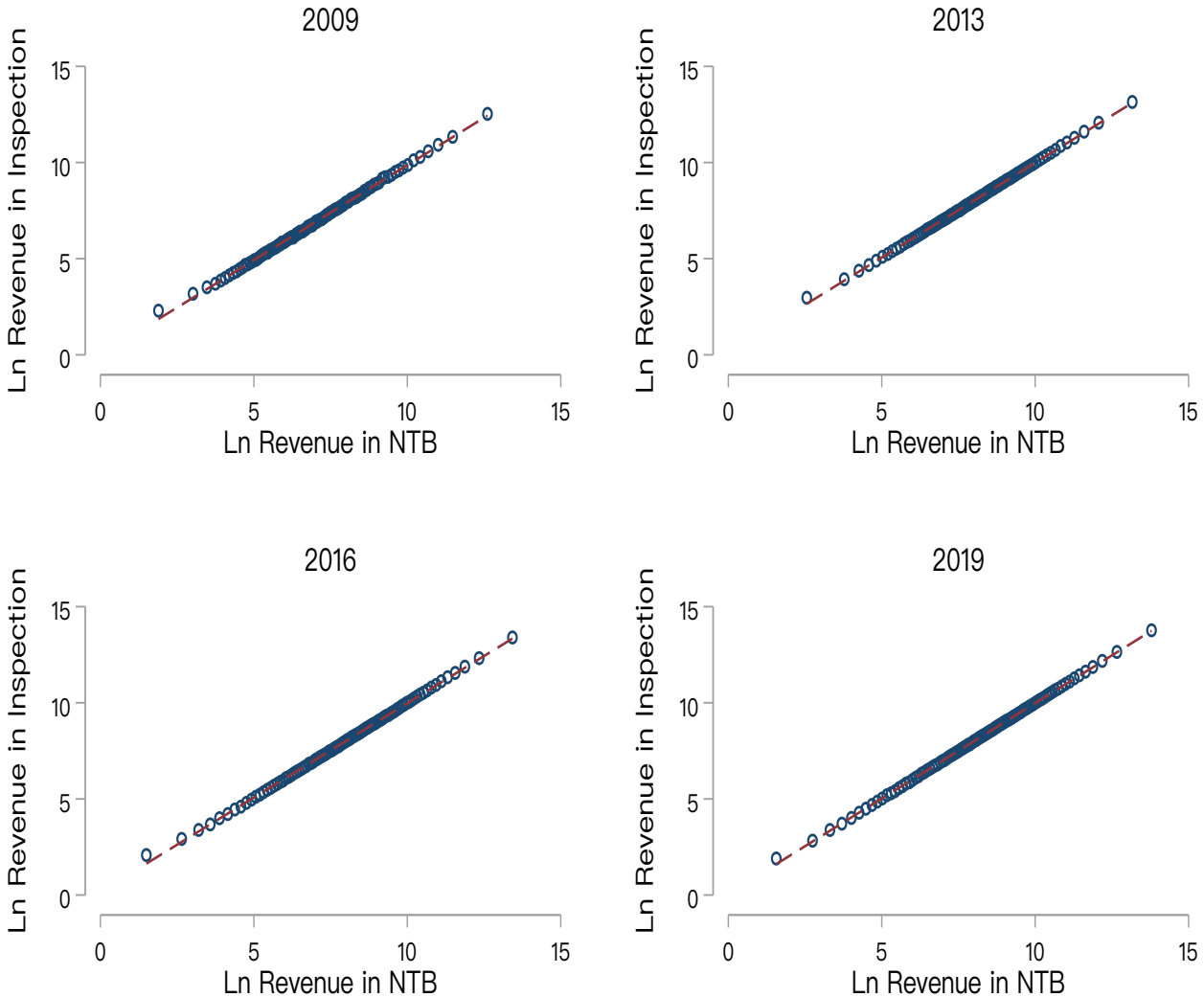








Bin Scatter Plot for Total Assets, Inspection vs. NTB



Sample is firms that are both in NTB data and Inspection Data (before imputation)

Adjustments to the Inspection Data

- Drop a class of firms that are missing in the NTB tax data (Sole proprietorship companies, partnerships)
- Remove fake firms
- Adjust date of entry and exit
- Impute missing data using K-Nearest Neighbor algorithm
- Ensure reporting is for standalone firm (not *group* level)

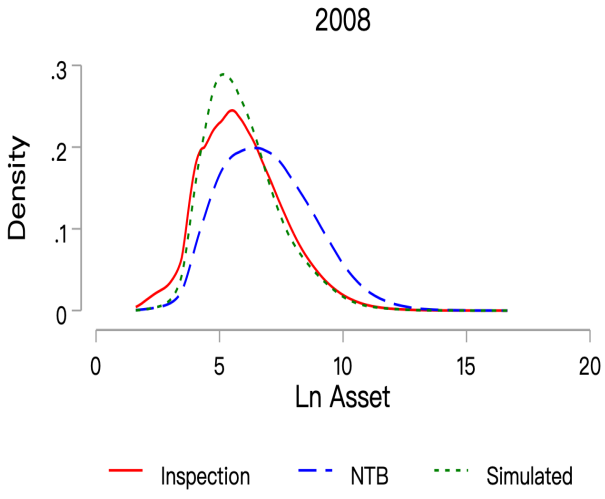
Inspection and NTB Sample Comparison

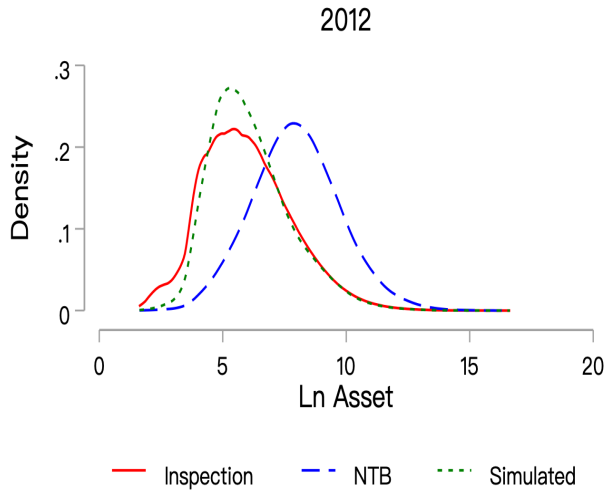
	Inspection data			NTB data			
Year	Number of firms	Frac. above scale	Frac. newly shown in data	Number of firms	Frac. above scale	Frac. newly shown in data	Frac. matched with registry
2008	1,254,414	38.7	11.6	242,603	60.8	31.2	93.9
2010	1,378,496	39.4	10.3	255,910	70.8	28.7	95.6
2012	1,590,326	16.2	10.3	226,242	57.0	18.6	96.3
2014	1,800,261	14.9	8.6	159,154	62.0	30.2	96.8
2016	1,973,742	13.3	8.5	163,788	58.1	12.8	97.2
2019	2,244,015	13.9	9.8	159,817	61.0	7.2	94.2

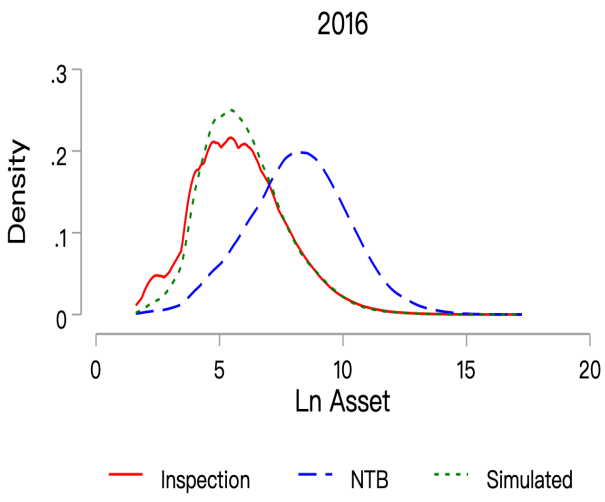
Weighting Strategy (Hellerstein and Imbens, 1999)

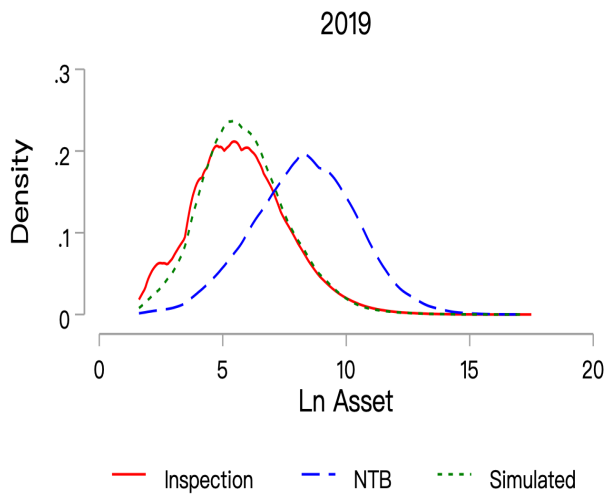
- Sort Inspection Data firms into bins on year, size, age, & industry
- Estimate true density in each “cell” from Inspection Data
- For each cell, generate weights for observations in NTB survey that are representative of their weight in the target population (Inspection Data, the universe)
- Problematic to separately estimate two densities and (non-parametrically) take ratio
- Instead: estimate density ratio by Least Square Importance Fitting

Distributions of (ln) Asset

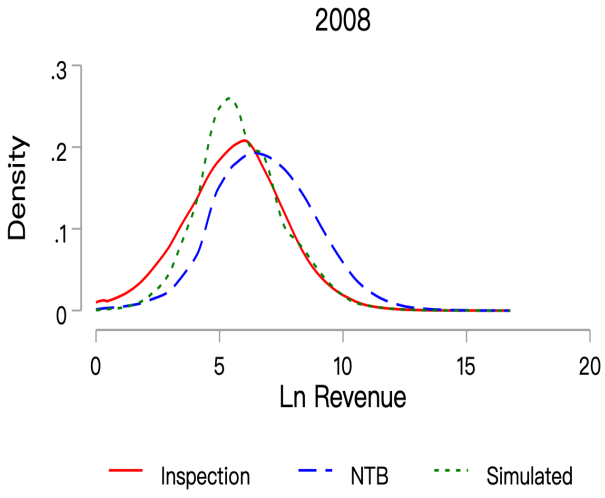


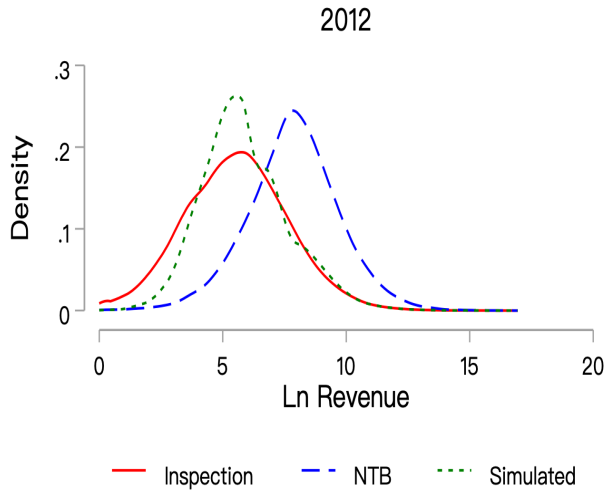


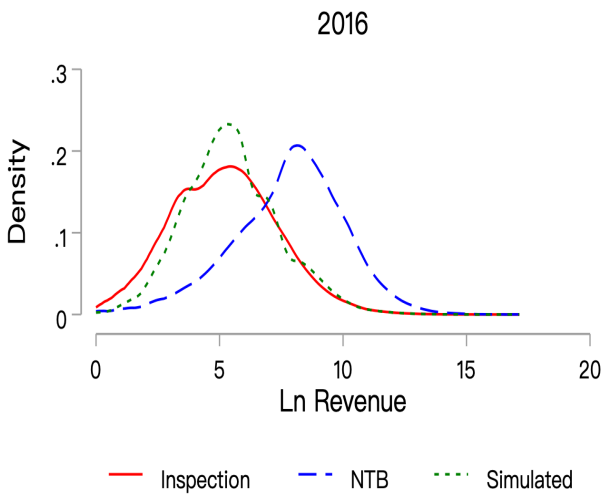


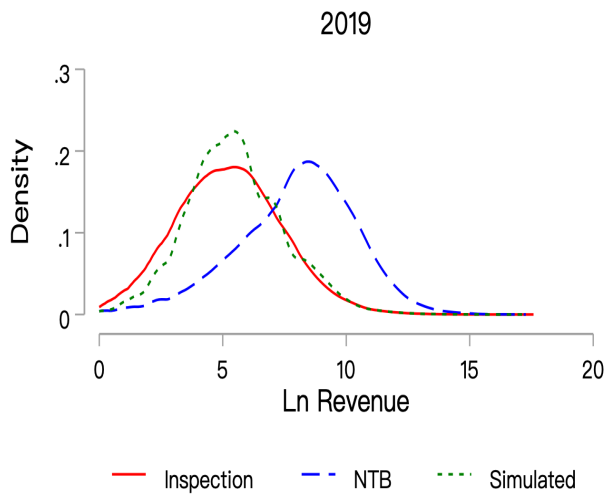


Distributions of (ln) Revenue

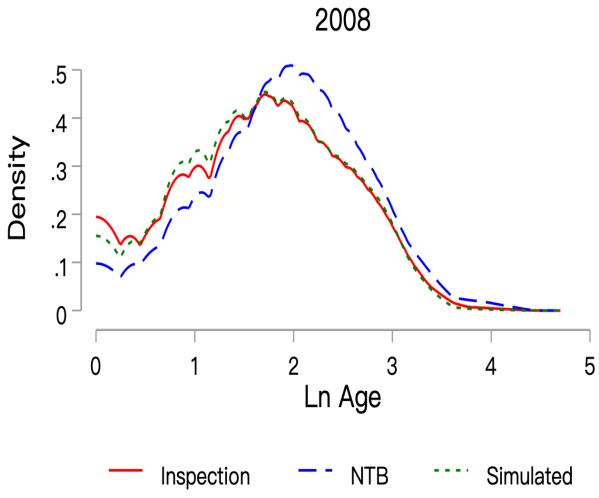


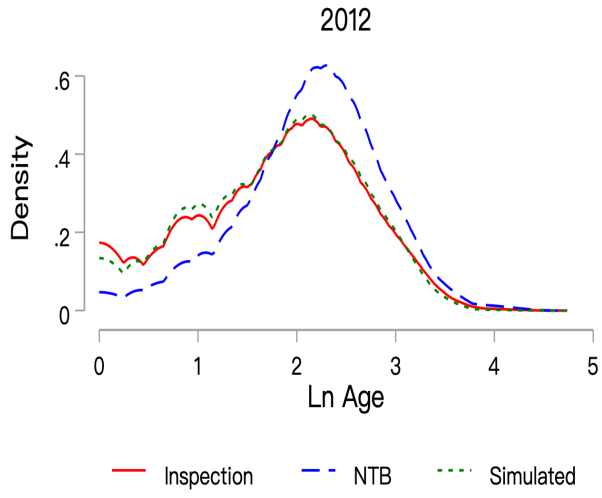


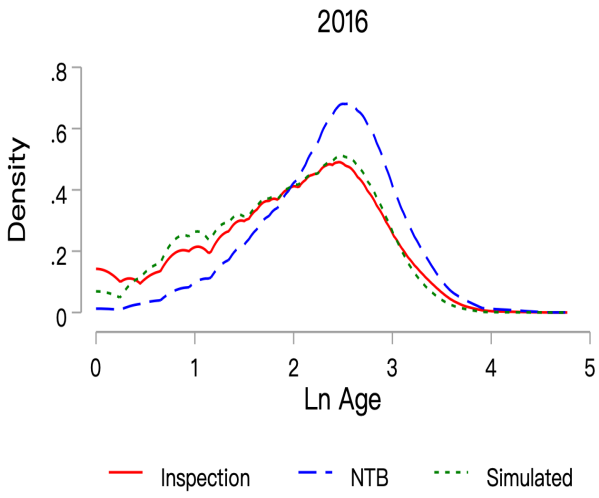


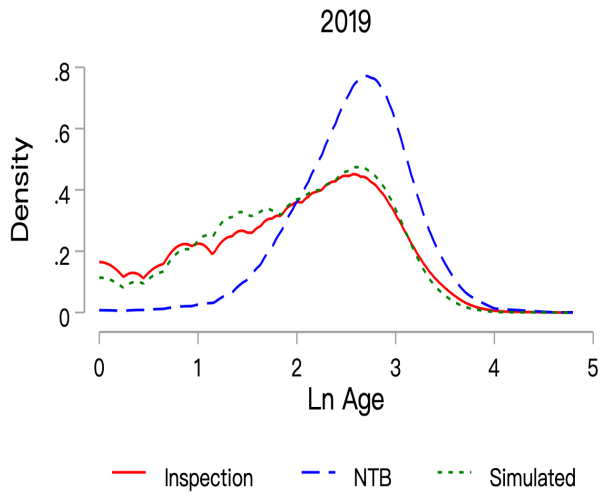


Distributions of (ln) Age









Manufacturing: Annualized Growth Rates of Key Variables

	Value added (nominal)	Gross output (nominal)	Employment (persons)	Tot. assets (nominal)	Wage per worker (nom.)
1998-2007					
NBS above-scale survey	21.1	20.7	4.6	12.9	12.0
2007-2019					
Inspection Data (rev. & assets)	-	7.1	-	8.9	-
NTB (weighted)	8.0	7.3	-1.5	9.2	11.3
2007-2011					
Inspection Data (rev. & assets)	-	13.3	-	14.6	-
NTB (weighted)	12.2	13.7	1.0	15.2	14.4
2011-2019					
Inspection Data (rev. & assets)	-	3.9	-	6.0	-
NTB (weighted)	5.9	4.1	-2.8	6.2	9.7

- Labor-income share not constant: fall 1998-2007, increase 2007-2019
- Value added as share of gross output increasing 2011-2019

Next steps

- Analyze growth in aggregate productivity (Solow residual)
- Quantify efficient growth vs. misallocation improvement
- Identify drivers of growth
  – forensic exploration of churn, size, industries, etc.

Theory: CES Production Function

- Assume:
  1. CES production fxn (allows varying labor-income share)
  2. Decreasing returns to scale $\eta < 1$
  3. Common labor-augmenting technical change λ_t

$$y_{it} = z_{it} \cdot \left[\left((k_{it})^{\frac{\rho-1}{\rho}} + (\lambda_t n_{it})^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \right]^{\eta}.$$

- Decreasing returns to scale: define aggregate TFP as

$$Z_t = \frac{Y_t}{(M_t)^{1-\eta} \left[X_t (K_t, N_t) \right]^{\eta}}$$

where

$$X(k,n) = \left(k^{\frac{\rho-1}{\rho}} + (\lambda n)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

Motivate Definition of Aggregate TFP

- Consider special case **without** heterogeneity & wedges $\Rightarrow$ all firms have same allocation. Value added per firm:

$$y_i = Z \cdot [X(k_i, n_i)]^{\eta} = Z \cdot \left[X\left(\frac{K}{M}, \frac{N}{M}\right) \right]^{\eta}$$

- Z is the natural measure of aggregate TFP
- Aggregate output is $Y = M \cdot y_i$ so

$$Z = \frac{\frac{Y}{M}}{\left[X\left(\frac{K}{M}, \frac{N}{M}\right) \right]^{\eta}} = \frac{Y}{M^{1-\eta} [X(K, N)]^{\eta}}$$

Misallocation

- Firms face output and capital frictions (τ_i, τ_{ik})
- Distorted problem is

$$\max_{k_i, n_i} \{(1-\tau_i) y_i - w n_i - (1+\tau_{ik})(r+\delta) k_i\}$$

- Note: definition of aggr. TFP incorporates distortions

Efficient Allocation and Aggregate TFP

- Analyze efficient allocation (to assess role of distortions)
- Efficient allocation: $y_{it}^* = (z_{it})^{1/(1-\eta)} \cdot \bar{y}_t^*$ where

$$\bar{y}_t^* = \left[(K_t)^{\frac{\rho-1}{\rho}} + (\lambda_t N_t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\eta\rho}{\rho-1}} \left(\int (z_i)^{\frac{1}{1-\eta}} \, dF_t(i) \right)^{-\eta}.$$

- Aggregate efficient TFP (same formula as for aggr. TFP)

$$\begin{aligned} Z_t^* &= \frac{\int y_i^* \, dF_t(i)}{(M_t)^{1-\eta} \left[\left((K_t)^{\frac{\rho-1}{\rho}} + (\lambda_t N_t)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \right]^{\eta}} \\ &= \left(\frac{1}{M_t} \int (z_i)^{\frac{1}{1-\eta}} \, dF_t(i) \right)^{1-\eta}. \end{aligned}$$

Identify λ_t Under Distortions

- Calibrate $\rho = 0.6$ (Yang-Zhang 2024)
  and $\eta = 2/3$ (Hsieh-Klenow 2009)
- Want to measure λ_t
  Problem: in the presence of wedges, aggregate data no longer identifies the sequence $\{\lambda_t\}_{t=1}^{T}$
- Solution: make two assumptions about output wedges,
  1. In the initial period, capital and output wedges have zero contribution to the aggregate labor-income share (normalization of λ_1).
  2. After 6 years of age, output wedges are constant over firm's life cycle

Identify λ_t (continued)

- Distributions F_t and F_{t+1} denote representative samples of incumbent firms operating in both t and $t+1$ ("synthetic cohort")
- Firm-level output wedge given by FOC for labor:

$$1-\tau_i = \frac{1}{\eta}\frac{w_t n_{ti}}{y_{ti}}\left(\left(\frac{k_{ti}}{\lambda_t n_{ti}}\right)^{\frac{\rho-1}{\rho}}+1\right).$$

- Integrate this equation over incumbent firms in periods $t+1$ and t,

$$\underbrace{\frac{\int \eta\,(1-\tau_i)\,dF_{t+1}(i)}{\int \eta\,(1-\tau_i)\,dF_t(i)}}_{=1} = \frac{\int\left[\left(\left(\frac{n_{i,t+1}}{k_{i,t+1}}\right)^{\frac{1-\rho}{\rho}}(\lambda_{t+1})^{\frac{1-\rho}{\rho}}+1\right)\frac{n_{i,t+1}w_{t+1}}{y_{i,t+1}}\right]dF_{t+1}(i)}{\int\left[\left(\left(\frac{n_{it}}{k_{it}}\right)^{\frac{1-\rho}{\rho}}(\lambda_t)^{\frac{1-\rho}{\rho}}+1\right)\frac{n_{it}w_t}{y_{it}}\right]dF_t(i)}$$

- Identifies sequence of $\{\lambda_t\}$ given λ_1

Hicks-Neutral Firm-Level TFP z_{it}

- Measure firm-level TFP z_{it} by inverting production function

$$z_{it} = y_{it} \cdot \left(\left((k_{it})^{\frac{\rho-1}{\rho}} + (\lambda_t n_{it})^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \right)^{-\eta}.$$

Annualized Growth Rates in λ, Z, Z^*, M, K, N, and Y

	Year	λ_g	Z_g	Z_g^*	M	K	N	Y
NBS	1998-2007	17.5	5.6	5.4	11.6	10.4	4.9	19.1
NTB	2007-19	9.0	1.1	3.5	4.7	4.5	-1.5	6.6
	2007-10	5.3	4.6	6.7	5.4	7.3	0.2	10.7
	2010-19	10.2	0.0	2.4	4.5	3.5	-2.1	5.3
Account ΔY_g		-	36%	15.2%	18.4%	-	-	100%

Growth Accounting

$$\begin{aligned}\Delta \ln Y_t &\approx \Delta \ln Z_t + (1-\eta)\cdot \Delta \ln M_t + \eta s_t \cdot \Delta \ln K_t \\ &\quad + \eta(1-s_{t-1})\cdot[\Delta \ln \lambda_t + \Delta \ln N_t] \\ s_{t-1} &= \left[1+(\lambda_{t-1}N_{t-1}/K_{t-1})^{(\rho-1)/\rho}\right]^{-1}.\end{aligned}$$

	Year	Z_g	λ_g	M_g	K_g	N_g
NBS	1998-2007	29.2	21.1	20.2	23.8	5.9
NTB	2007-19	17.3	44.2	23.8	23.0	-7.6
	2007-10	43.5	16.7	16.7	22.5	0.6
	2010-19	-0.4	62.5	28.6	23.0	-12.9

Straw Man: Standard Productivity Decomposition

- What is role of **churn** (entry, ext, realloc.) in accounting for $\{Z_t, Z_t^*\}$?
- Standard BHC-FHK is not useful because formulae not intended to study aggregate Solow residual and misallocation
- Assume $Z_t = \sum_i s_{it} z_{it}$ for some weights s_{it}.

  Baily–Hulten–Campbell decomp. of aggr. productivity growth:

$$\Delta Z \approx \underbrace{\sum_{i \in C} \bar{s}_{it} \Delta z_{it}}_{\text{within}} + \underbrace{\sum_{i \in C} (\bar{z}_{it} - \bar{Z}_t) \Delta s_{it}}_{\text{between}} + \underbrace{\sum_{i \in E} s_{it} (z_{it} - \bar{Z}_t)}_{\text{entry}} - \underbrace{\sum_{i \in X} s_{i,t-1} (z_{i,t-1} - \bar{Z}_t)}_{\text{exit}}$$

  where

$$\bar{s}_{it} = (s_{it} + s_{i,t-1})/2 \qquad \bar{z}_{it} = (z_{it} + z_{i,t-1})/2 \qquad \bar{Z}_t = (Z_t + Z_{t-1})/2$$

Straw Man (cont.)

- In our setting,

$$Z = \frac{Y}{[X(K,N)]^{\eta}} = \sum_i z_i \cdot \left[\underbrace{X(k_i,n_i)/X(K,N)}_{\text{input share}} \right]^{\eta}$$

- Problems with BHC-FHK decomp.
  - s_i would be "share" (of inputs) only if $\eta = 1$ (unrealistic)
  - within term $\neq$ efficient growth $\rightarrow$ silent about misallocation
  - ignores interaction between entry and incumbents

Aggregate TFP: Additive Group Decomposition

- Define aggr. TFP for group G, Z_G, and input weight ω_G

$$
\begin{aligned}
Z_{Gt} &\equiv \frac{\sum_{i \in G} Y_{it}}{(M_{Gt})^{1-\eta} \left[X_t(K_{Gt}, N_{Gt})\right]^{\eta}} \\
\omega_{Gt} &\equiv \left(\frac{M_{Gt}}{M_t}\right)^{1-\eta} \left(\frac{X_t(K_{Gt}, N_{Gt})}{X_t(K_t, N_t)}\right)^{\eta}
\end{aligned}
$$

- TFP contributions are additive,

$$
Z_t = \sum_G \underbrace{\omega_{Gt}}_{\text{input "weight"}} \cdot \underbrace{Z_{Gt}}_{\text{TFP of group } G} .
$$

Aggregation Result: Additive Decomp. of Z^*

- Define efficient TFP and efficient input weight for group G

$$Z_{Gt}^* \equiv \frac{\sum_{i \in G} y_{it}^*}{(M_{Gt})^{1-\eta} \left[X_t\left(K_{Gt}^*, N_{Gt}^*\right)\right]^{\eta}} = \left(\frac{1}{M_{Gt}} \int_G (z_i)^{\frac{1}{1-\eta}} \, dF_t(i)\right)^{1-\eta}$$

$$\omega_{Gt}^* \equiv \left(\frac{M_{Gt}}{M_t}\right)^{1-\eta} \left(\frac{X_t\left(K_{Gt}^*, N_{Gt}^*\right)}{X_t\left(K_t, N_t\right)}\right)^{\eta} = \left(\frac{Z_{Gt}^*}{Z_t^*}\right)^{\frac{\eta}{1-\eta}} \frac{M_{Gt}}{M_t}$$

- Result: efficient TFP is additive in Z_G^*

$$Z_t^* = \sum_G \omega_{Gt}^* \cdot Z_{Gt}^*$$

Realized Growth vs. Efficient Growth

Realized TFP growth is efficient growth plus reduction in misallocation,

$$\underbrace{\ln\left(\frac{Z_t}{Z_{t-1}}\right)}_{\text{realized growth}} = \underbrace{\ln\left(\frac{Z_t^*}{Z_{t-1}^*}\right)}_{\text{efficient growth}} + \underbrace{\ln\left(\frac{Y_t}{Y_{t-1}}\right) - \ln\left(\frac{Y_t^*}{Y_{t-1}^*}\right)}_{\text{improvement in misallocation}}$$

(Shapley) Decomposition of TFP Growth from Churn

$$\ln\left(\frac{Z_t}{Z_{t-1}}\right) \approx \underbrace{\bar{\omega}_{Ct}\frac{\Delta Z_{C,t}}{\bar{Z}_t}}_{\text{Incumbent growth}} + \underbrace{\left(\frac{\bar{Z}_{Ct}}{\bar{Z}_t}-1\right)\Delta\omega_{C,t}}_{\text{Reallocation to incumbents}} + \underbrace{\omega_{Nt}\left(\frac{Z_{N,t}}{\bar{Z}_t}-1\right)}_{\text{Entry}} + \underbrace{\omega_{X,t-1}\left(1-\frac{Z_{X,t-1}}{\bar{Z}_t}\right)}_{\text{Exit}}$$

where

$$\Delta Z_{Ct} \equiv Z_{C,t}-Z_{C,t-1} \qquad \Delta\omega_{Ct} \equiv \omega_{Ct}-\omega_{C,t-1}$$

$$\bar{Z}_{Ct} \equiv \frac{Z_{C,t}+Z_{C,t-1}}{2} \qquad \bar{\omega}_{Ct} \equiv \frac{\omega_{Ct}+\omega_{C,t-1}}{2}$$

Decomp. of TFP Growth: Incumbents, Entry, Exit

$$\underbrace{\ln\left(\frac{Z_t}{Z_{t-1}}\right)}_{\substack{0\\ \text{aggr. growth}}} \approx \underbrace{\bar{\omega}_C \frac{\Delta Z_{C,t}}{\bar{Z}}}_{\substack{1\\ \text{incumb. grwth}}} + \underbrace{\left(\frac{\bar{Z}_C}{\bar{Z}} - 1\right)\Delta\omega_{Ct}}_{\substack{2\\ \text{realloc. to incumb.}}} + \underbrace{\omega_{Nt}\left(\frac{Z_{Nt}}{\bar{Z}} - 1\right)}_{\substack{3\\ \text{entry}}} + \underbrace{\omega_{\xi,t-1}\left(1 - \frac{Z_{\xi,t-1}}{\bar{Z}}\right)}_{\substack{4\\ \text{exit}}}$$

Year	0 Total	1 Incumb.	2 Realloc.	3 Entry	4 Exit
1998-2007	0.056	0.069	-0.002	-0.024	0.021
2007-2019	0.011	0.028	0.000	-0.013	0.002
2007-2010	0.046	0.066	-0.001	-0.015	0.003
2010-2019	0.000	0.016	0.000	-0.012	0.002

TFP Growth: Efficient vs. Misallocation

- Incumbent Growth

$$\bar{\omega}_{Ct}\frac{\Delta Z_{Ct}}{\bar{Z}_t} = \underbrace{\bar{\omega}^*_{Ct}\frac{\Delta Z^*_{C,t}}{\bar{Z}^*_t}}_{\text{efficient growth}} + \underbrace{\bar{\omega}_{Ct}\frac{\Delta Z_{Ct}}{\bar{Z}_t} - \bar{\omega}^*_{Ct}\frac{\Delta Z^*_{Ct}}{\bar{Z}^*_t}}_{\text{reduction in misallocation}}$$

- Reallocation to Incumbents

$$\left(\frac{\bar{Z}_{Ct}}{\bar{Z}_t} - 1\right)\Delta\omega_{Ct} = \underbrace{\left(\frac{\bar{Z}^*_{Ct}}{\bar{Z}^*_t} - 1\right)\Delta\omega^*_{Ct}}_{\text{efficient growth}}$$

$$+ \underbrace{\left(\frac{\bar{Z}_{Ct}}{\bar{Z}_t} - 1\right)\Delta\omega_{Ct} - \left(\frac{\bar{Z}^*_{Ct}}{\bar{Z}^*_t} - 1\right)\Delta\omega^*_{Ct}}_{\text{reduction in misallocation}}$$

TFP Growth: Efficient vs. Misallocation

- Entry

$$\omega_{N,t}\left(\frac{Z_{N,t}}{\bar{Z}_t}-1\right)=\underbrace{\omega^*_{Nt}\left(\frac{Z^*_{Nt}}{\bar{Z}^*_t}-1\right)}_{\text{efficient growth}}+\underbrace{\omega_{Nt}\left(\frac{Z_{Nt}}{\bar{Z}_t}-1\right)-\omega^*_{Nt}\left(\frac{Z^*_{Nt}}{\bar{Z}^*_t}-1\right)}_{\text{reduction in misallocation}}$$

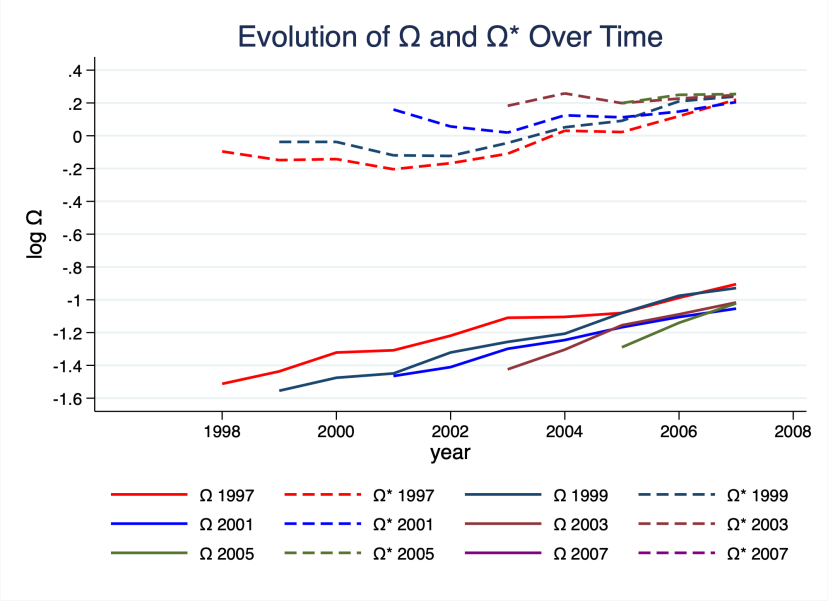
- Exit

$$\begin{aligned}\omega_{X,t-1}\left(1-\frac{Z_{X,t-1}}{\bar{Z}_t}\right) &= \underbrace{\omega^*_{X,t-1}\left(1-\frac{Z^*_{X,t-1}}{\bar{Z}^*_t}\right)}_{\text{efficient growth}}\\ &+ \underbrace{\omega_{X,t-1}\left(1-\frac{Z_{X,t-1}}{\bar{Z}_t}\right)-\omega^*_{X,t-1}\left(1-\frac{Z^*_{X,t-1}}{\bar{Z}^*_t}\right)}_{\text{reduction in misallocation}}\end{aligned}$$

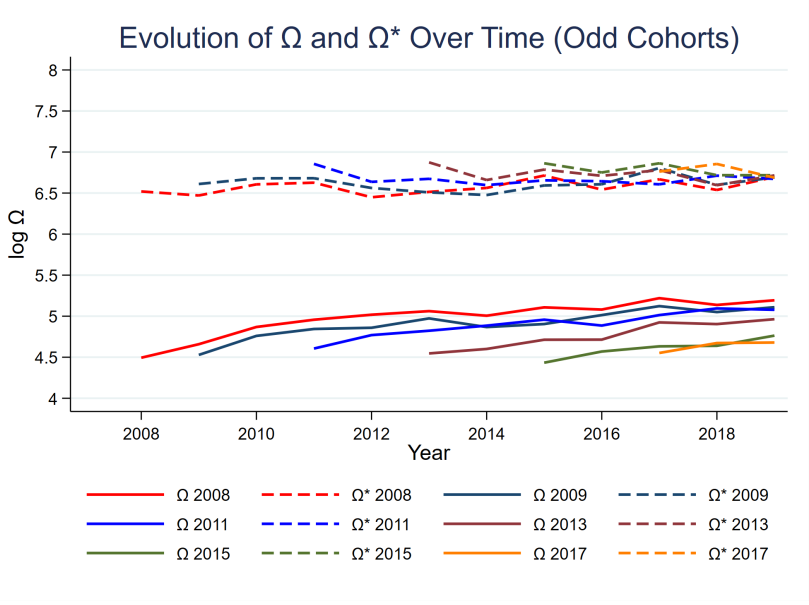
Decomposition of TFP Growth

	Incumbent		Reallocation		Entry		Exit	
Year	Eff.	Misall.	Eff.	Misall.	Eff.	Misall.	Eff.	Misall.
1998-07	2.1	4.9	0.4	-0.7	4.2	-6.6	-1.0	3.1
2007-19	1.8	1.0	0.1	-0.1	16.0	-17.3	-7.5	7.8
2007-10	4.4	2.1	0.1	-0.1	17.4	-19.0	-6.8	7.1
2010-19	0.9	0.6	0.1	-0.1	15.6	-16.8	-7.8	8.0

Z and Z^*, by Cohort

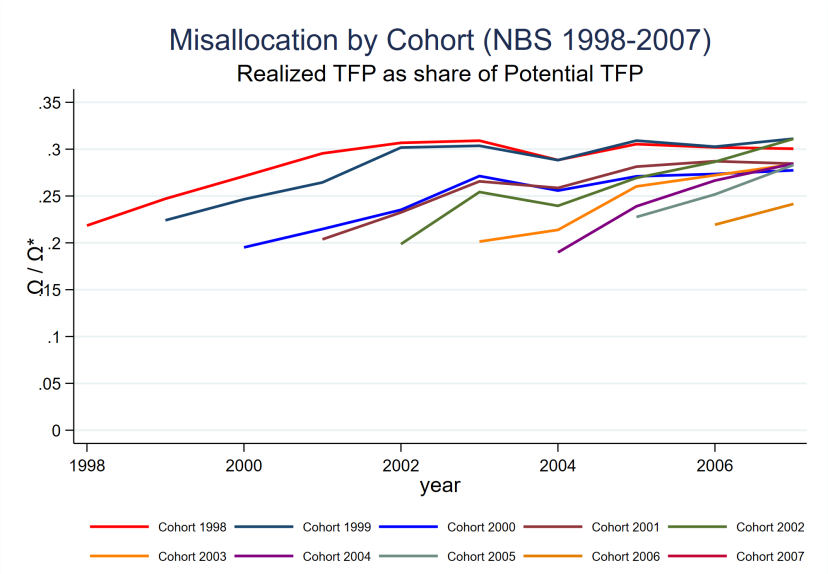


(a) NBS: 1998–2007

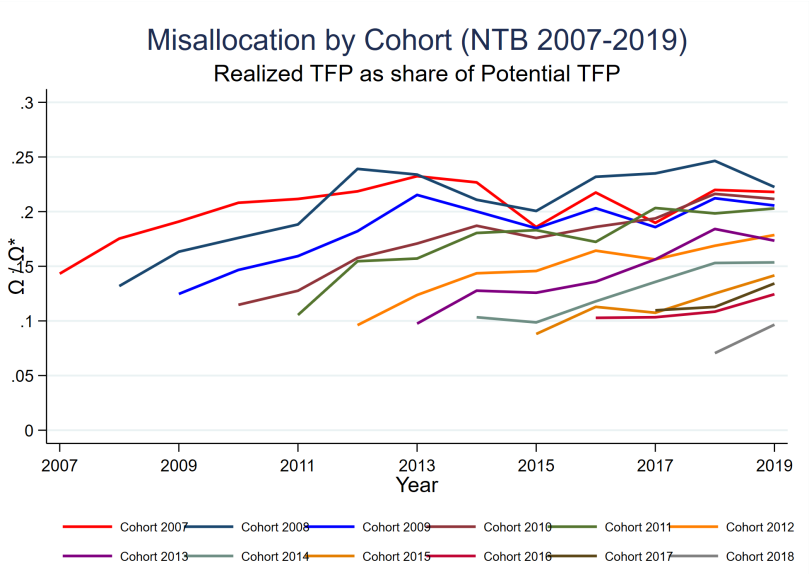


(b) NTB: 2007–2019

Misallocation over the Life Cycle Z/Z^*



(a) NBS: 1998–2007. Measured TFP as share of potential TFP



(b) NTB: 2007–2019. Measured TFP as share of potential TFP

Summary

- New firms have lower measured TFP than incumbents
  ... but higher efficient TFP
- **Before 2007**:
  - Misallocation falls over life cycle
  - New firms catch up with incumbents' measured TFP
  - Incumbents catch up with new firms' efficient TFP
  - High incumbent growth due to rapid reduction in misallocation for **young** incumbent firms
- **After 2011**: new firms' misallocation lingers over life cycle
- Needed: firm-dynamics theory
  - generating positive selection of new firms' TFP
  - ... and depressed measured TFP for new firms
- Solution: firm-level distortions, dissipating over life cycle

General Shapley Decomposition of Growth

- Divide firm into mutually exclusive groups
  - examples: size, industry, incumbent/entry/exit

$$\ln\left(\frac{Z_t}{Z_{t-1}}\right) = \underbrace{\sum_G \bar{\omega}_{Gt}\frac{\Delta Z_{G,t}}{\bar{Z}_t}}_{\text{average TFP growth}} + \underbrace{\sum_G \Delta\omega_{G,t}\frac{\bar{Z}_{Gt}}{2\bar{Z}_t}}_{\text{reallocation}}$$

$$\bar{\omega}_{Gt} = \frac{\omega_{G,t-1}+\omega_{Gt}}{2} \qquad \Delta\omega_{G,t} = \omega_{G,t}-\omega_{G,t-1}$$

$$\bar{Z}_{Gt} = \frac{Z_{G,t-1}+Z_{Gt}}{2} \qquad \Delta Z_{G,t} = Z_{G,t}-Z_{G,t-1}$$

$$\bar{Z}_t = \frac{Z_{t-1}+Z_t}{2}$$

- Decompose realized TFP growth into efficient growth versus change in misallocation,

$$
\begin{aligned}
\ln\left(\frac{Z_t}{Z_{t-1}}\right) &= \ln\left(\frac{Z_t^*}{Z_{t-1}^*}\right) + \underbrace{\ln\left(\frac{Y_t}{Y_{t-1}}\right) - \ln\left(\frac{Y_t^*}{Y_{t-1}^*}\right)}_{\text{misallocation}} \\
&\approx \underbrace{\sum_G \bar{\omega}_{G,t}^* \frac{\Delta Z_{G,t}^*}{\bar{Z}_t^*}}_{\text{average eff. TFP growth}} + \underbrace{\sum_G \Delta\omega_{Gt}^* \frac{\bar{Z}_{Gt}^*}{\bar{Z}_t^*}}_{\text{efficient reallocation}} \\
&+ \underbrace{\sum_G \bar{\omega}_{G,t} \frac{\Delta Z_{G,t}}{\bar{Z}_t} - \bar{\omega}_{G,t}^* \frac{\Delta Z_{G,t}^*}{\bar{Z}_t^*}}_{\text{misallocation improvement.: avg. TFP growth}} \\
&+ \underbrace{\sum_G \left[\Delta\omega_{Gt} \frac{\bar{Z}_{Gt}}{\bar{Z}_t} - \Delta\omega_{Gt}^* \frac{\bar{Z}_{Gt}^*}{\bar{Z}_t^*}\right]}_{\text{misallocation improvement.: reallocation}}
\end{aligned}
$$

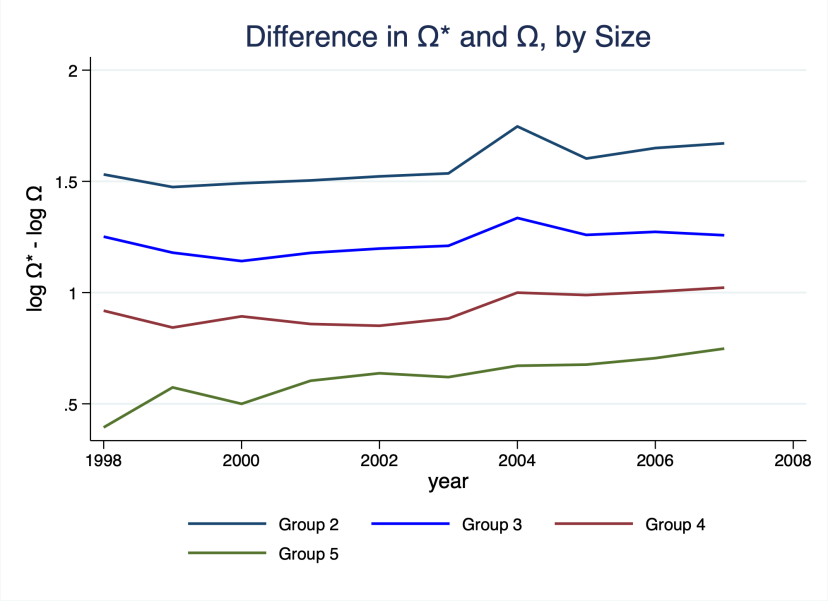
Decomposition of TFP Growth, by Size Groups

	Year	Z^*	Z	TFP average	TFP realloc.	Misall. average	Misall. realloc.
NBS							
	1998-07	5.4	5.6	3.9	1.5	-0.9	1.2
NTB							
	2007-19	3.5	1.1	3.4	0.1	-3.6	1.3
	2007-10	6.7	4.6	5.8	0.9	-4.3	2.3
	2010-19	2.4	0.0	2.5	-0.2	-3.4	1.0

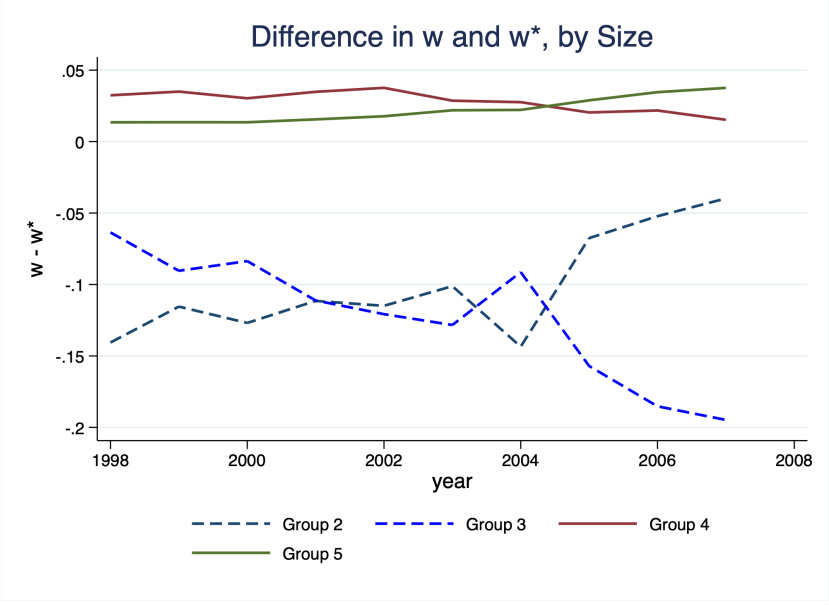
Decomposition of TFP Growth, by Size Groups

		Z^*	Z	Efficient TFP		Misallocation	
				average	reallocation	average	reallocation
NBS: 1998–07		5.4	5.6	3.9	1.5	-0.9	1.2
	2	-1.4	-0.6	0.5	-1.9	-0.4	1.1
Size group	3	4.2	1.9	2.6	1.6	-0.7	-1.6
	4	2.0	2.1	0.6	1.4	0.2	0.0
	5	0.5	2.2	0.1	0.4	0.0	1.7
NTB: 2007–19		3.5	1.1	3.4	0.1	-3.6	1.3
	1	0.7	-0.1	0.6	0.1	-0.7	-0.1
	2	1.8	-0.2	1.4	0.4	-1.4	-0.6
Size group	3	0.9	-0.5	1.1	-0.2	-1.0	-0.4
	4	-0.3	0.5	0.1	-0.4	-0.4	1.2
	5	0.3	1.4	0.1	0.2	-0.1	1.3

TFP Growth, NBS 1998-2007: by Size Groups

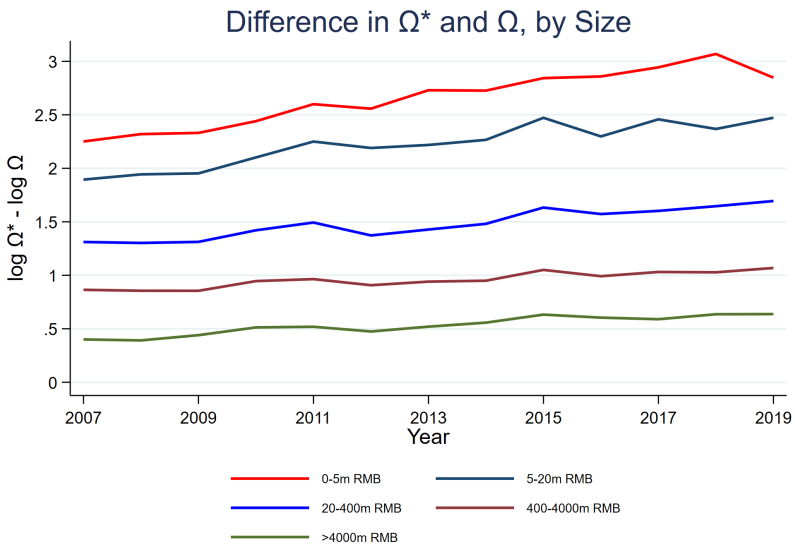


(a) Z and Z^*

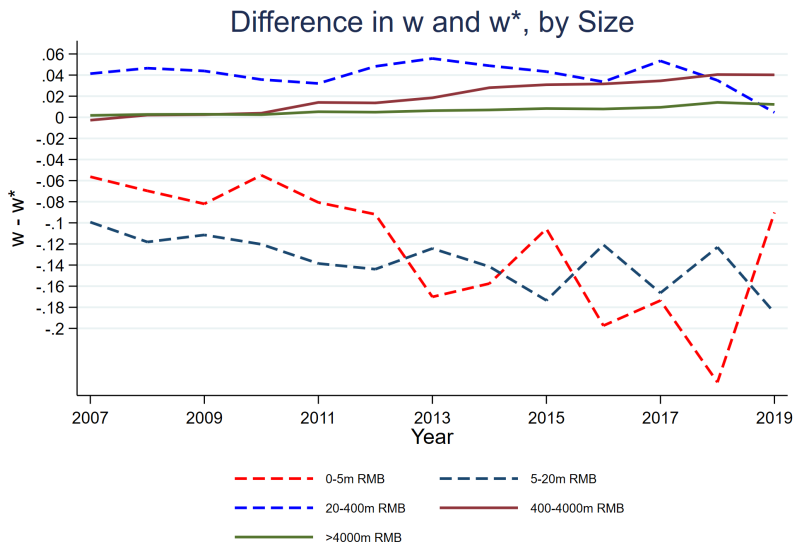


(b) ω and ω^*

TFP Growth, NTB 2007-2019: by Size Groups



(a) Z and Z^*



(b) ω and ω^*

Decomposition of TFP Growth, by Regions

		Z^*	Z	TFP		Misallocation	
				within	between	within	between
NBS: 1998–07		5.4	5.6	5.4	0.0	0.0	0.2
	C	0.5	0.5	0.6	-0.1	0.2	-0.2
	E	3.4	3.1	2.7	0.7	-0.2	-0.1
	N	0.4	0.5	0.8	-0.4	0.1	0.1
Region	NE	0.5	0.4	0.3	0.2	0.1	-0.2
	NW	0.1	0.1	0.1	0.0	0.0	0.0
	S	0.1	0.7	0.6	-0.5	0.0	0.6
	SW	0.6	0.3	0.4	0.2	0.0	-0.3
NTB: 2007–19		3.5	1.1	3.4	0.1	-3.6	1.3
	C	-0.4	0.0	0.2	-0.6	0.0	0.5
	E	1.3	0.7	1.6	-0.3	-0.8	0.3
	N	0.2	0.0	0.4	-0.2	-0.2	0.1
Region	NE	0.7	0.0	0.3	0.4	-0.4	-0.3
	NW	0.1	0.1	0.1	0.0	-0.1	0.0
	S	1.4	0.2	0.7	0.7	-0.6	-0.6
	SW	0.4	0.1	0.2	0.2	-0.2	-0.1

Model

- Dynamic Hopenhayn model (endogenous entry/exit)
- Age-dependent distortions (dispersion falling with age)
- Account for new firms being more selected (higher firm-level TFP) but lower aggregate TFP (more distorted)
- High incumbent growth due to reduction in misallocation (for young firms) and exit (for older and more misallocated firms)
- Model elements: general TFP growth, entry costs, distribution of new firms relative to existing firms, financial constraints, size-dependent policies
- Estimate model using documented facts as moments

Conclusion

- New representative data set for manufacturing firms in China
- Calculate firm-level TFP and decomposition of aggr. TFP growth
- Large fall in TFP growth and labor-augmenting growth
- Half of growth fall due to lower incumb. avg. TFP growth, half due to falling role of churn (entry/exit)
- Pre-2007 growth for incumbents driven by reduction in misallocation over life cycle for young firms
- Post-2007 fall in TFP growth: misallocation dissipates more slowly over life cycle

Population of Interest (Inspection Data)

1. Sample Selection criteria

- **Legal Types:** We exclude *Getihu*, Partnerships, and Proprietorships due to negligible size and poor data quality (Exclusion rate: ∼4–9%).
- **New Entrants:** We drop firms in their first year of establishment to ensure consistent financial reporting.
- **Industries:** Manufacturing, excluding Tobacco (16), Petroleum& Coal Mfg (25), and Other Mfg (41). (Exclusion rate: ∼6–9%).

2. Entity Classification (Final Sample) Based on registration type, we classify remaining firms into four groups:

1. **Unincorp:** Non-company legal persons (e.g., Traditional SOEs, Collectives).
2. **Domestic LLC:** Limited Liability Companies with domestic ownership.
3. **Foreign LLC:** LLCs with Foreign, HK, Macau, or Taiwan investment.
4. **Shareholding:** Corporations limited by shares (Public and Non-public).

Claim: NTB & Inspection Data Have High Quality

- Match firms in NTB and Inspection Data on Legal Repr. ID, firm name, and address (match $>$ 95% of NTB firms)
- Evidence of data quality:
  Assume classical m.e. in revenue & assets $\Rightarrow$ identification of m.e.:

$$\begin{aligned} \ln(y_{NTB}) &= ln(y) + \varepsilon_{NTB} \\ \ln(y_{ID}) &= ln(y) + \varepsilon_{ID} \\ \Rightarrow \quad var(\varepsilon_{NTB}) &= var\left[\ln(y_{NTB})\right] - cov\left[\ln(y_{NTB}, \ln(y_{ID})\right] \end{aligned}$$

Year	Variance of ln(Revenue)		Covariance between NTB and Ins.	Measurement Error in NTB	Measurement Error in Ins.
	NTB	Ins.			
2008	1.969	2.157	1.978	-0.009	0.179
2010	2.064	2.170	2.062	0.003	0.108
2012	1.986	2.075	1.979	0.007	0.096
2014	2.198	2.320	2.200	-0.002	0.120
2016	2.264	2.358	2.261	0.004	0.097
2019	2.521	2.581	2.519	0.002	0.062

Table: Revenue

Inspection vs. NTB: Classical M.E. for Assets

Year	Variance of ln(Asset)		Covariance between NTB and Ins.	Measurement Error in NTB	Measurement Error in Ins.
	NTB	Ins.			
2008	2.459	2.687	2.482	-0.023	0.205
2010	2.546	2.690	2.544	0.002	0.146
2012	2.364	2.501	2.374	-0.010	0.127
2014	2.728	2.872	2.740	-0.012	0.132
2016	2.697	2.821	2.707	-0.010	0.114
2019	2.838	2.929	2.846	-0.008	0.083

Table: Assets

- Sample is firms that are both in NTB data and Inspection Data (before imputation)

Weighting & Sampling: NTB $\rightarrow$ Inspection Alignment

Step 1: Stratification: Year $\times$ Size category $\times$ Age bin $\times$ Industry

- Size categories (NTB rules): 0–5m, 5–20m, 20–400m, 400-4000m, $\geq$4 billion
- Age bins: 1–4, 5–8, 9–12, 12+ years
- Industry: 10 aggregated manufacturing sectors

Step 2: Within-stratum weight estimation: estimate density ratio (weighting) function using:

- $\ln$(assets), $\ln$(revenue), $\ln$(age)
- Ownership type, Region
- Exit dummy (within 3 years)

Predict a *weight* for each NTB observation $\rightarrow$ aligns within-stratum covariate distribution to Inspection.

Step 3: Draw sample

- Deterministic retention of all NTB observations
- Supplementary bootstrap for remaining required count

Within-Strata Weighting Function Estimation Least-Square Fitting by Kanamori et al. (2009)

Let $\mathbf{v}$ denote the covariates used in weighting. For each stratum, let $f^T(\mathbf{v})$ and $f^S(\mathbf{v})$ be the joint distribution of $\mathbf{v}$ in the target (Inspection) and source (NTB) samples, respectively.

- The true density-ratio is $r^*(\mathbf{v}) = \frac{f^{T*}(\mathbf{v})}{f^{S*}(\mathbf{v})}$
- Estimate $r(\mathbf{v})$ minimizes the following mean squared error

$$\begin{aligned} SQ(r) &:= \frac{1}{2}\int (r(\mathbf{v}) - r^*(\mathbf{v}))^2 f^{S*}(\mathbf{v})d\mathbf{v} \\ &= \frac{1}{2}\int r(\mathbf{v})^2 f^{S*}(\mathbf{v})d\mathbf{v} - \int r(\mathbf{v}) f^{T*}(\mathbf{v})d\mathbf{v} + \frac{1}{2}r^*(\mathbf{v})^2 f^{S*}(\mathbf{v})d\mathbf{v} \end{aligned}$$

- The empirical counterpart for the first two terms is

$$\widehat{SQ}(r) = \frac{1}{2n^S}\sum_{j=1}^{n^S} r(\mathbf{v}_j^S)^2 - \frac{1}{n^T}\sum_{i}^{n^T} r(\mathbf{v}_i^T)$$

Algorithm

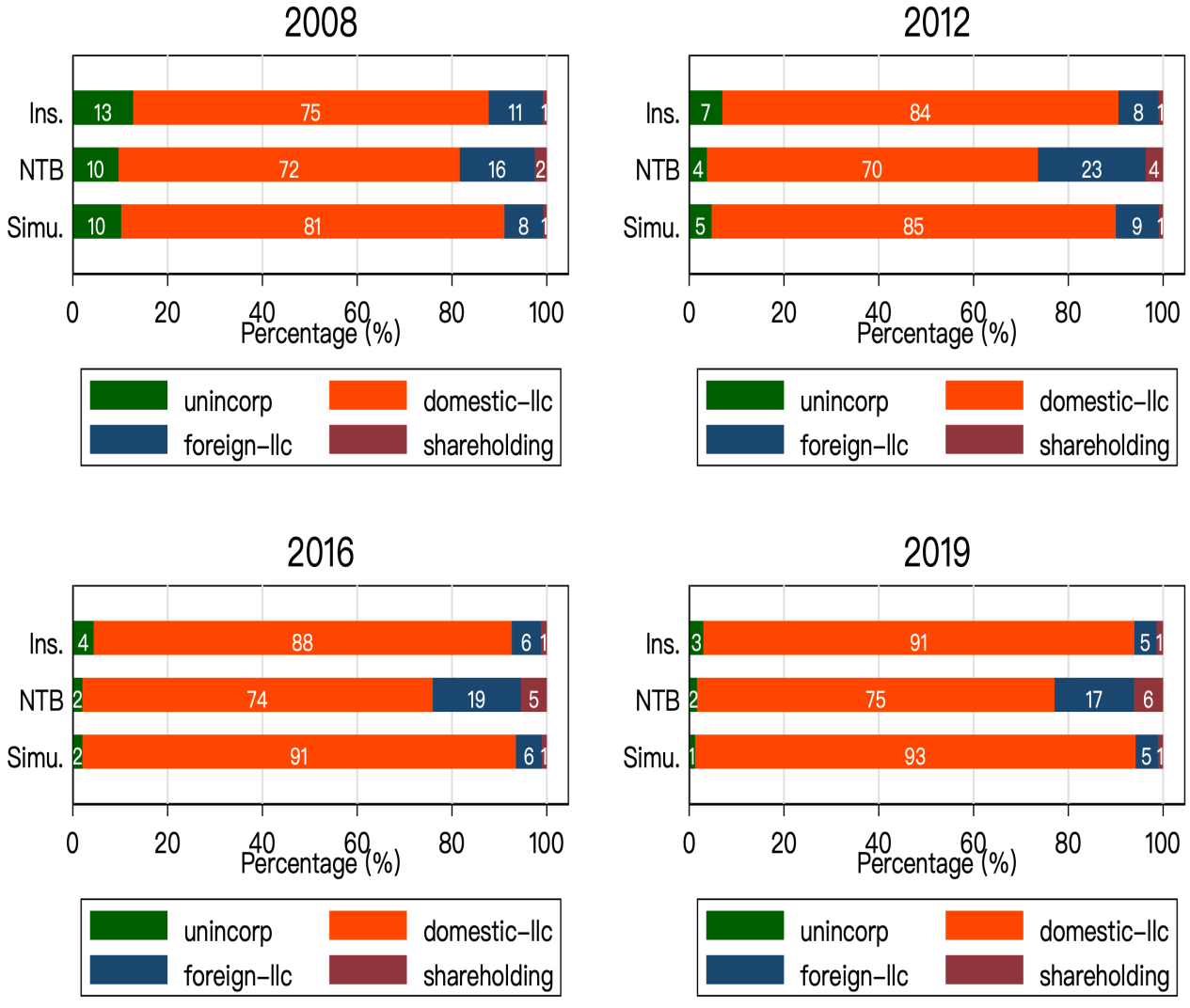
Approximate $r(\mathbf{v})$ with linear combination of basis functions $\sum_{c=1}^{C} \alpha_c \phi_c(\mathbf{v})$

- Take C kernel centers from data
- $\{\phi_c(\mathbf{v})\}_{c=1}^{C}$ are radial basis functions for distance of $\mathbf{v}$ to each kernel center c
- $\alpha' = (\alpha_1, \alpha_2, ..., \alpha_C)$ are the combination weights to be estimated

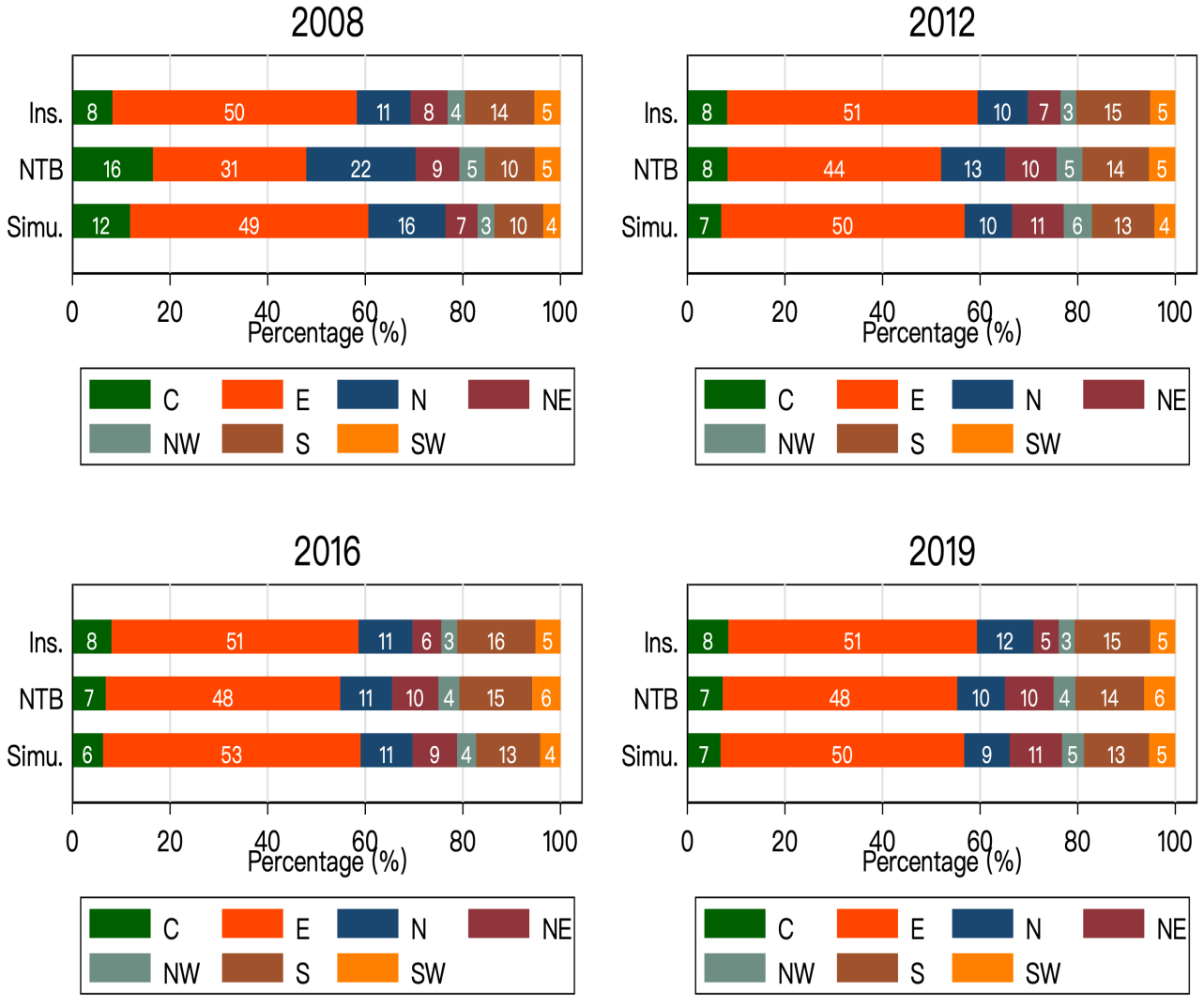
The objective function becomes

$$\widehat{SQ}(\alpha) = \frac{1}{2} \sum_{c=1}^{C} \sum_{c'=1}^{C} \alpha_c \alpha_{c'} \left(\frac{1}{n^S} \sum_{i=1}^{n^S} \phi_c(\mathbf{v}_i^S) \phi_{c'}(\mathbf{v}_i^S) \right) - \sum_{c=1}^{C} \alpha_c \left(\frac{1}{n^T} \sum_{j=1}^{n^T} \phi_c(\mathbf{v}_j^T) \right)$$

Distributions of Entity Type (Ownership)



Distributions of Location



Large Changes in Aggr. Labor-Income Share

	Year	Labor-income share $\frac{wN}{Y}$
NBS	1998	31.0%
	2007	20.5 %
NTB	2007	35.5%
	2019	43.5%

Reason for difference between NTB and NBS in 2007:

- Simulated sample includes below-scale firms whereas NBS has only above-scale firms
- Calculate value added in NTB using income approach

Why We Need CES

- Suppose zero output wedges and no heterogeneity in capital wedges across firms. Assume constant ρ and η
- Implied growth in λ_t and Z_t

$$\Delta \ln \lambda_t = \Delta \ln \left(\frac{K_t}{N_t} \right) + \frac{\rho}{1-\rho} \cdot \Delta \ln \left(\frac{\eta Y_t}{w_t N_t} - 1 \right)$$

$$\begin{aligned} \Delta \ln Z_t = {} & \Delta \ln Y_t - (1-\eta) \Delta \ln (M_t) - \eta \Delta \ln (K_t) \\ & - \frac{\rho \eta}{1-\rho} \Delta \ln \left(1 - \frac{w_t N_t}{\eta Y_t} \right) \end{aligned}$$

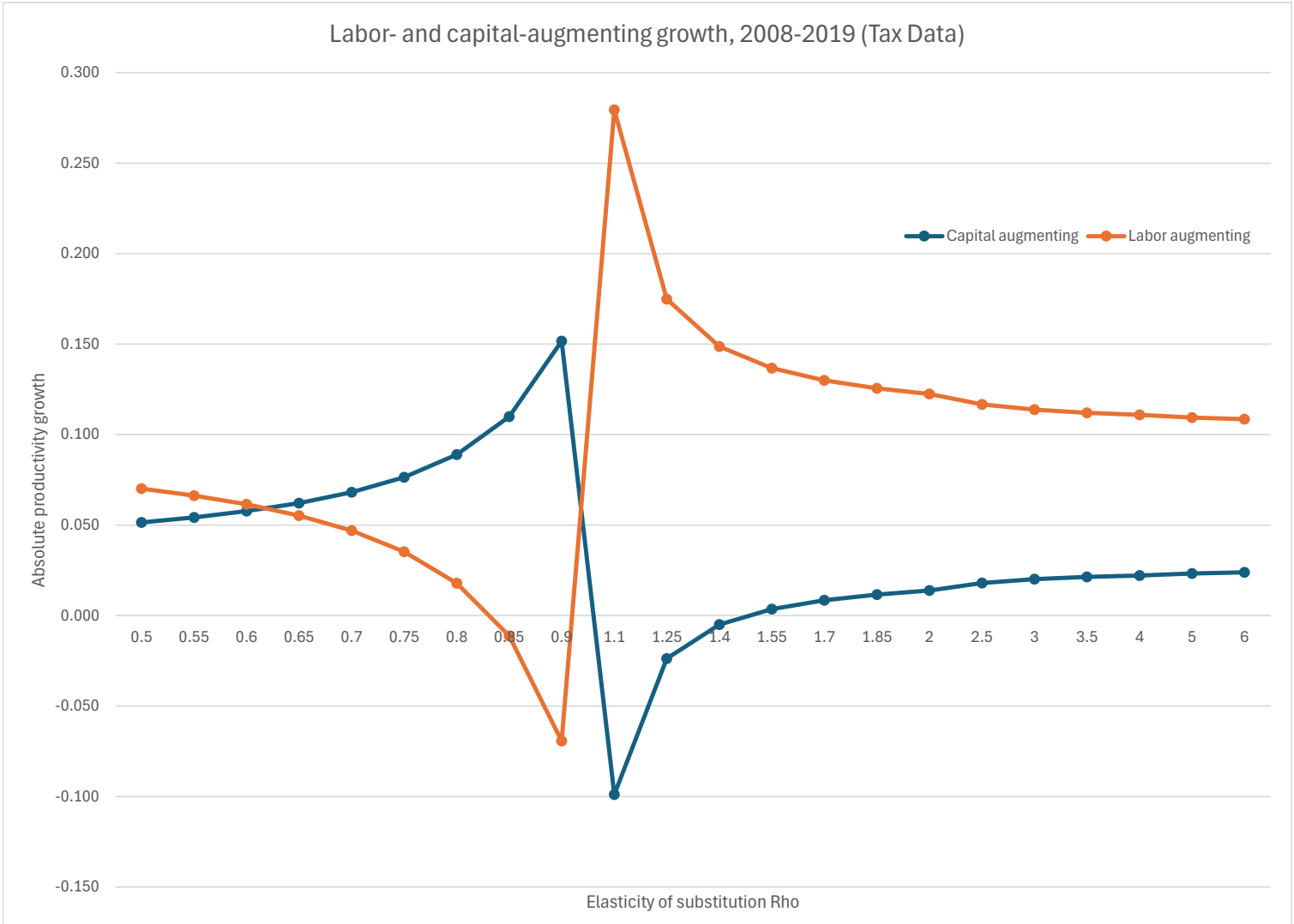
- Diamond, McFadden, and Rodriguez (1978): ρ cannot be separately identified from λ_t

(Hicks-neutral) Growth Accounting (cont.)

- Fix $\eta = 2/3$ (inconsequential)
- Calculate implied rates in Z_t and λ_t for different ρ
- Plausible theoretical restriction: absolute labor- and capital-augmenting growth should not be negative (both Z_t and $Z_t\lambda_t$ should grow over time)

$$\frac{Y_t}{M_t} = \left(\left(\left(\underbrace{Z_t}_{\text{capital-aug.}} \frac{K_t}{M_t} \right)^{\frac{\rho-1}{\rho}} + \left(\underbrace{Z_t\lambda_t}_{\text{labor-aug.}} \frac{N_t}{M_t} \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \right)^{-\eta} (Z_t)^{1-\eta}$$

2008-2019 Implied Factor-Augm. Growth



1998-2007 Implied Factor-Augm. Growth



What Drives the Aggregate Labor-Income Share?

Decomposition of **growth** in aggregate labor-income share

$$\ln \frac{\frac{w_t N_t}{Y_t}}{\frac{w_1 N_1}{Y_1}} = \underbrace{\left[\left(\frac{1}{\lambda_1}\frac{K_1}{N_1}\right)^{\frac{1-\rho}{\rho}} + 1\right]^{-1}\left(1 - \left(\frac{\frac{K_t}{K_1}}{\frac{\lambda_t}{\lambda_1}\cdot\frac{N_t}{N_1}}\right)^{\frac{\rho-1}{\rho}}\right)}_{\text{contribution of capital deepening \& labor-augmenting growth}}$$

$+$ contribution of output wedges

$+$ contribution of capital wedges

What Drives the Aggregate Labor-Income Share?

- Suppose output wedges and TFP are jointly log-normal

$$\ln \frac{\frac{w_t N_t}{Y_t}}{\frac{w_1 N_1}{Y_1}} \approx \underbrace{\left[\left(\frac{1}{\lambda_1}\frac{K_1}{N_1}\right)^{\frac{1-\rho}{\rho}} + 1\right]^{-1}\left(1 - \left(\frac{\frac{K_t}{K_1}}{\frac{\lambda_t}{\lambda_1}\cdot\frac{N_t}{N_1}}\right)^{\frac{\rho-1}{\rho}}\right)}_{\text{contribution of capital deepening \& labor-augmenting growth}}$$

$$+ \underbrace{\Delta E\{\ln(1-\tau)\} + \frac{1+\eta}{1-\eta}\frac{1}{2}\Delta var\left[\ln(1-\tau)\right] + \frac{1}{1-\eta}\Delta cov\left[\ln(1-\tau), \ln(z)\right]}_{\text{contr. of output wedges}}$$

$$+ \text{contribution of capital wedges}$$

Accounting for Changes in Labor-Income Shares

		Growth in $\log \frac{wN}{Y}$	Capital growth	Labor-augmenting	Output wedge	Capital wedge
NBS	1998-2007	-0.34	+0.15	-0.45	-0.20	+0.19
NTB	2008-2014	+0.06	+0.14	-0.25	-0.04	+0.21
	2008-2019	+0.13	-	-	-	-