Lecture 4: Distortions and Misallocation

Kunal Sangani

ECON 416-1

Recap: Aggregation Results for Efficient Economies

- Solow (1957) for economies with aggregate production functions:

$$d \log Y = d \log A + \sum_f \Lambda_f d \log L_f.$$

- Hulten (1978) for disaggregated economies:

$$d \log A = \sum_i \lambda_i d \log A_i, \quad \text{where} \quad \lambda_i = \frac{p_i y_i}{\sum_k p_k y_k}.$$

- Nonlinearities depend on change in sales shares. E.g., CES economy with one factor:

$$\frac{d\lambda_i}{d \log A_i} = \sum_{j=0}^{N} (\theta_j - 1) \lambda_j Var_{\Omega^{(j)}}(\Psi_{(i)}).$$

The Solow residual in a simple economy

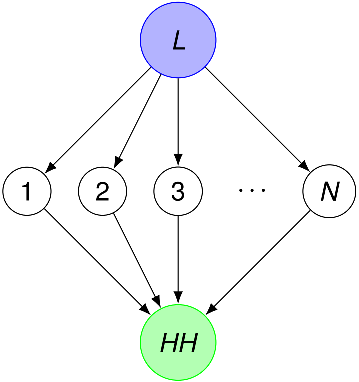


Figure: Horizontal economy.

- Consider a horizontal economy with N firms,

$$y_i = A_i L_i,$$

$$Y = \left(\sum_{i=1}^{N} y_i^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}},$$

$$L = \sum_{i=1}^{N} L_i.$$

- We can write,

$$Y = \underbrace{\left(\sum_{i=1}^{N} \left(A_i \frac{L_i}{L} \right)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}}_{A = \text{Solow residual}} L.$$

The Solow residual in a simple economy

- Denote share of labor used each firm by $l_i = L_i / L$.

$$Y = AL = \left(\sum_{i=1}^{N} (A_i l_i)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}} L.$$

- Clearly, Y depends both on technological primitives $(\{A_i\}, L)$ and how resources are allocated across firms $\{l_i\}$.

- We can decompose changes in output into changes in each of these components.

$$d \log Y = \underbrace{\sum_i \frac{\partial \log Y}{\partial \log A_i} + d \log L}_{\text{Changes in technology}} + \underbrace{\sum_i \frac{\partial \log Y}{\partial \log l_i} d \log l_i}_{\text{Reallocations across firms}} .$$

- Suppose we increase a firm's productivity? Which of these would change?

The Solow residual in a simple economy

- Consider a perturbation dl_i. Feasibility requires $\sum_i dl_i = 0$. Can we improve output?

$$d \log A = \sum_i \frac{(A_i l_i)^{\frac{\theta-1}{\theta}}}{\sum_{k=1}^{N} (A_k l_k)^{\frac{\theta-1}{\theta}}} d \log l_i$$

The Solow residual in a simple economy

- Consider a perturbation dl_i. Feasibility requires $\sum_i dl_i = 0$. Can we improve output?

$$
\begin{aligned}
d \log A &= \sum_i \frac{(A_i l_i)^{\frac{\theta-1}{\theta}}}{\sum_{k=1}^N (A_k l_k)^{\frac{\theta-1}{\theta}}} d \log l_i \\
&= \sum_i \frac{\lambda_i}{l_i} dl_i
\end{aligned}
$$

The Solow residual in a simple economy

- Consider a perturbation dl_i. Feasibility requires $\sum_i dl_i = 0$. Can we improve output?

$$
\begin{aligned}
d\log A &= \sum_i \frac{\left(A_i l_i\right)^{\frac{\theta-1}{\theta}}}{\sum_{k=1}^{N}\left(A_k l_k\right)^{\frac{\theta-1}{\theta}}} d\log l_i \\
&= \sum_i \frac{\lambda_i}{l_i} dl_i \\
&= Cov\left(\frac{\lambda_i}{l_i}, dl_i\right)
\end{aligned}
$$

The Solow residual in a simple economy

- Consider a perturbation dl_i. Feasibility requires $\sum_i dl_i = 0$. Can we improve output?

$$
\begin{aligned}
d\log A &= \sum_i \frac{(A_i l_i)^{\frac{\theta-1}{\theta}}}{\sum_{k=1}^{N} (A_k l_k)^{\frac{\theta-1}{\theta}}} d\log l_i \\
&= \sum_i \frac{\lambda_i}{l_i} dl_i \\
&= Cov\left(\frac{\lambda_i}{l_i}, dl_i\right)
\end{aligned}
$$

- In efficient economy, sales and cost shares coincide. $\Rightarrow Cov(1, dl_i) = 0$.
- Intuition: Efficient economy equates marginal value of resource across all uses.

$$
\frac{\partial \log Y}{\partial L_i} = \frac{\partial \log Y}{\partial \log L_i} \frac{1}{L_i} = \frac{\lambda_i}{L_i}.
$$

- Distortions (taxes/markups, quotas, central planner, etc.) can cause this to fail.

Table of Contents

Hsieh and Klenow (2009): Setup

- Canonical model of misallocation from Hsieh and Klenow (2009).
  - See also Banerjee and Duflo (2005) survey on motivating micro evidence.

- Output is a Cobb-Douglas aggregate across sectors $s = 1, ..., S$,

$$Y = \prod_{s=1}^{S} Y_s^{\theta}, \qquad \text{where} \qquad \sum_{s=1}^{S} \theta_s = 1.$$

- Each sector's output is CES aggregate of M_s firms' products,

$$Y_s = \left(\sum_{i=1}^{M_s} y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}.$$

- Each firm produces with capital and labor,

$$y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}.$$

  Assumption: capital and labor shares identical within sector.

Hsieh and Klenow (2009): Setup

- Output distortion and capital distortion for each firm, so that profits are

$$\pi_{si} = (1-\tau_{Ysi})p_{si}y_{si} - wL_{si} - (1+\tau_{Ksi})rK_{si}.$$

  - τ_{Ysi} high: e.g., high taxes, government restrictions on size, transportation frictions.
  - τ_{Ksi} high: e.g., costly access to credit.

- Profit-maximizing prices

$$p_{si} = \frac{\sigma}{\sigma-1}\frac{(1+\tau_{Ksi})^{\alpha_s}}{(1-\tau_{Ysi})}\frac{1}{A_{si}}\left(\frac{r}{\alpha_s}\right)^{\alpha_s}\left(\frac{w}{1-\alpha_s}\right)^{1-\alpha_s}.$$

- Capital-labor ratio and output across firms are

$$\frac{K_{si}}{L_{si}} = \frac{\alpha_s}{1-\alpha_s}\frac{w}{r}\frac{1}{1+\tau_{Ksi}}, \qquad Y_{si} \propto \left(A_{si}\frac{(1-\tau_{Ysi})}{(1+\tau_{Ksi})^{\alpha_s}}\right)^{\sigma}.$$

Marginal revenue products

- Marginal revenue product of labor

$$MRPL_{si} = p_{si}\frac{\partial y_{si}}{\partial L_{si}} = (1-\alpha_s)\frac{p_{si}y_{si}}{L_{si}} = \frac{\sigma}{\sigma-1}w\frac{1}{1-\tau_{Ysi}}.$$

- Marginal revenue product of capital

$$MRPK_{si} = p_{si}\frac{\partial y_{si}}{\partial K_{si}} = \alpha_s\frac{p_{si}y_{si}}{K_{si}} = \frac{\sigma}{\sigma-1}r\frac{1+\tau_{Ksi}}{1-\tau_{Ysi}}.$$

- In the absence of wedges τ_{Ksi}, τ_{Ysi}, MRPL and MRPK equalized across firms.

Output and TFPR

- Output in each sector

$$
\begin{aligned}
Y_s = \left(\sum_i y_{si}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} &= \left(\sum_i \left(A_{si}\left(\frac{K_{si}}{L_{si}}\right)^{\alpha_s} L_{si}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \\
&\propto \left(\sum_i \left(A_{si}\frac{1-\tau_{Ysi}}{(1+\tau_{Ksi})^{\alpha_s}}\right)^{\sigma-1}\right)^{\frac{1}{\sigma-1}} \\
\Rightarrow Y_s &= \left(\sum_i \left(A_{si}\frac{\overline{TFPR}_s}{TFPR_{si}}\right)^{\sigma-1}\right)^{\frac{1}{\sigma-1}} K_s^{\alpha_s} L_s^{1-\alpha_s}.
\end{aligned}
$$

where $TFPR_{si} = p_{si}A_{si} \propto (MRPK_{si})^{\alpha_s}(MRPL_{si})^{1-\alpha_s}$, $\overline{TFPR}_s$ is the geometric average across firms.

Output and TFPR

$$Y_s = \left(\sum_i \left(A_{si} \frac{\overline{TFPR_s}}{TFPR_{si}} \right)^{\sigma-1} \right)^{\frac{1}{\sigma-1}} K_s^{\alpha_s} L_s^{1-\alpha_s}.$$

- If *TFPR* equal across all firms, $A_s = \left(\sum_i (A_{si})^{\sigma-1} \right)^{\frac{1}{\sigma-1}}$.

Output and TFPR

$$Y_s = \left(\sum_i \left(A_{si} \frac{\overline{TFPR_s}}{TFPR_{si}} \right)^{\sigma - 1} \right)^{\frac{1}{\sigma - 1}} K_s^{\alpha_s} L_s^{1-\alpha_s}.$$

- If *TFPR* equal across all firms, $A_s = \left(\sum_i (A_{si})^{\sigma-1} \right)^{\frac{1}{\sigma-1}}$.
- Suppose productivities A_i and $TFPR_{si}$ distributed according to joint lognormal. Then:

$$\log A_s = \underbrace{\log \left(\sum_i \left(A_{si} \frac{\overline{TFPR_s}}{TFPR_{si}} \right)^{\sigma - 1} \right)^{\frac{1}{\sigma - 1}}}_{\text{log } A_s \text{ at frontier}} - \underbrace{\frac{\sigma}{2} \mathrm{Var}(\log TFPR_{si})}_{\text{Distance to frontier}}.$$

Measurement

- We can infer wedges from capital-labor ratio and output-labor ratios:

$$1+\tau_{Ksi} = \frac{\alpha_s}{1-\alpha_s}\frac{wL_{si}}{rK_{si}},$$
$$1-\tau_{Ysi} = \frac{\sigma}{\sigma-1}\frac{1}{1-\alpha_s}\frac{wL_{si}}{p_{si}y_{si}}.$$

- Of course, this is coming from (strong) assumption of common Cobb-Douglas sectoral technologies.

Measurement

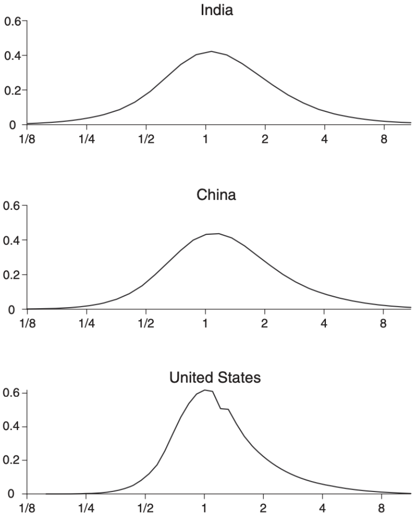


FIGURE II
Distribution of TFPR

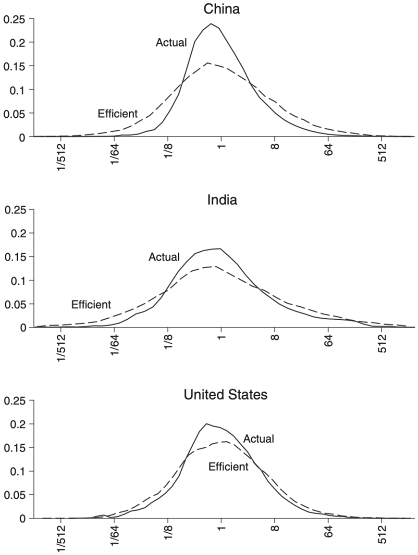


FIGURE III
Distribution of Plant Size

Gains

Year	Gains from moving to frontier	Gains from moving to 1997 US TFPR dispersion
China		
1998	115	51
2001	96	37
2005	87	31
India		
1987	100	40
1991	102	41
1994	128	59
United States		
1977	36	
1987	31	
1997	43	

- Manufacturing TFP in US 130% higher than China. $\Rightarrow$ Misallocation accounts for 49%.
- Manufacturing TFP in US 160% higher than India. $\Rightarrow$ Misallocation accounts for 35%.

Ownership and TFPR

TABLE VII
TFP BY OWNERSHIP

	TFPR	TFPQ
China		
State	−0.415	−0.144
	(0.023)	(0.090)
Collective	0.114	0.047
	(0.010)	(0.013)
Foreign	−0.129	0.228
	(0.024)	(0.040)
India		
State (central)	−0.285	0.717
	(0.082)	(0.295)
State (local)	−0.081	0.425
	(0.063)	(0.103)
Joint public/private	−0.162	0.671
	(0.037)	(0.085)

Notes. The dependent variable is the deviation of log TFPR or log TFPQ from the industry mean. The independent variables for China are dummies for state-owned plants, collective-owned plants (plants jointly owned by local governments and private parties), and foreign-owned plants. The omitted group is domestic private plants. The independent variables for India are dummies for a plant owned by the central government, a plant owned by a local government, and a plant jointly owned by the government (either central or local) and by private individuals. The omitted group is a privately owned plant (both domestic and foreign). Regressions are weighted least squares with industry value-added shares as weights. Entries are the dummy coefficients, with standard errors in parentheses. Results are pooled for all years.

- State-owned firms tend to have lower TFPR, consistent with better access to credit or preferential treatment.

Takeaways

- Influential proof-of-concept that misallocation can be large.
- Changes in allocative efficiency can matter.
- Reasonable to have concerns about strong assumptions needed to map to data.
- Today, we're going to relax many of the parametric assumptions (horizontal economy, lognormal distributions) to analyze the effect of any "wedges."

Table of Contents

Technical vs. Allocative Efficiency

- We can write output as a function of technology and resource allocation, $Y = \mathcal{Y}(\mathcal{A}, \mathcal{X})$.

- In our example, $\mathcal{A} = (\{A_i\}, L)$ and $\mathcal{X} = (\{l_i\})$.

- Then, we can always write:

$$d \log Y = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log \mathcal{A}} d \log \mathcal{A}}_{\Delta \text{Technological Efficiency}} + \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} d\mathcal{X}}_{\Delta \text{Allocative Efficiency}} .$$

- First term: Effect of technologies on output, holding fixed allocation of resources.

- Second term: Improvement in output due to reallocations across uses.

Technical vs. Allocative Efficiency: Distortions

- Let's introduce some set of distortionary instruments τ.
- E.g., in our horizontal economy ex, let $\tau = \{\tau_i\}$ be vector of taxes on each firm.
- Allocation of resources $\mathcal{X}$ will depend both on technologies and on taxes.
- We now can write output in terms of technology primitives and distortions:

$$Y = \mathcal{Y}(\mathcal{A}, \mathcal{X}(\mathcal{A}, \tau)).$$

- Technology change can have effects on *both* technical and allocative efficiency:

$$\frac{d \log Y}{d \log \mathcal{A}} = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log \mathcal{A}}}_{\Delta \text{Technological Efficiency}} + \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} \frac{\partial \mathcal{X}(\mathcal{A}, \tau)}{\partial \log A}}_{\Delta \text{Allocative Efficiency}}.$$

Technical vs. Allocative Efficiency: Efficient economy

- Suppose output is initially efficient given technologies $\mathcal{A}$. Then:

$$Y = \max_{\mathcal{X}} \mathcal{Y}(\mathcal{A}, \mathcal{X}) \qquad \Rightarrow \qquad \frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} = 0.$$

- By the envelope theorem,

$$\frac{d \log Y}{d \log \mathcal{A}} = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log \mathcal{A}}}_{\Delta\text{Technological Efficiency}} + \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} \frac{\partial \mathcal{X}(\mathcal{A}, \tau)}{\partial \log A}}_{\Delta\text{Allocative Efficiency}}.$$

Technical vs. Allocative Efficiency: Efficient economy

- Suppose output is initially efficient given technologies $\mathcal{A}$. Then:

$$Y = \max_{\mathcal{X}} \mathcal{Y}(\mathcal{A}, \mathcal{X}) \qquad \Rightarrow \qquad \frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} = 0.$$

- By the envelope theorem,

$$\frac{d \log Y}{d \log \mathcal{A}} = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log \mathcal{A}}}_{\Delta \text{Technological Efficiency}} + \underbrace{0}_{\Delta \text{Allocative Efficiency}}.$$

Technical vs. Allocative Efficiency: Efficient economy

- Suppose output is initially efficient given technologies $\mathcal{A}$. Then:

$$Y = \max_{\mathcal{X}} \mathcal{Y}(\mathcal{A}, \mathcal{X}) \qquad \Rightarrow \qquad \frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}} = 0.$$

- By the envelope theorem,

$$\frac{d \log Y}{d \log \mathcal{A}} = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log \mathcal{A}}}_{\Delta \text{Technological Efficiency}} + \underbrace{0}_{\Delta \text{Allocative Efficiency}}.$$

- Starting at efficiency, marginal reallocations have no effect on output.

- Even if endpoint is inefficient!

Table of Contents

Framework

- To characterize changes in allocative efficiency generally (i.e., away from initially efficient economies), we will put some structure on the problem.

- Arbitrary factors $f = 1, ..., F$, each in fixed supply L_f.

- Arbitrary producers $i = 1, ..., N$ with CRS production given by cost functions,

$$TC_i(y_i) = \frac{1}{A_i} C_i(p_{i1}, ..., p_{iN}, w_{i1}, ..., w_{iF}) y_i.$$

- Each producer N has an exogenous markup μ_i:

$$p_i = \mu_i \frac{1}{A_i} C_i.$$

- Remember: μ_i stand-in for any sorts of "wedges" (taxes, shadow price of quota, etc.)

Framework

- Final demand maximizes a homothetic aggregator,

$$Y = \max_{\{c_i\}} \mathcal{D}(c_1, ..., c_N),$$

  subject to budget constraint

$$\sum_i p_i c_i = \sum_f w_f L_f + \sum_i \pi_i,$$

  where profits π_i given by

$$\pi_i = \left(1 - \frac{1}{\mu_i}\right) p_i y_i.$$

- No rent seeking. (See Liu (2019) for an example where wedge revenue is wasted.)

Framework: Generality

1. Input-specific wedges.

2. Consumption taxes / wedges.

3. Endogenous wedges (e.g., endogenous markups).

Framework: Generality

1. Input-specific wedges.
    - $TC_i(y_i) = \frac{1}{A_i} C_i((1+\tau_{i1})p_1, ..., (1+\tau_{iN})p_N, (1+\tau^f_{i1})w_1, ..., (1+\tau^f_{iF})w_F) y_i.$
2. Consumption taxes / wedges.
3. Endogenous wedges (e.g., endogenous markups).

Framework: Generality

1. Input-specific wedges.
   - $TC_i(y_i) = \frac{1}{A_i} C_i((1+\tau_{i1})p_1, ..., (1+\tau_{iN})p_N, (1+\tau_{i1}^f)w_1, ..., (1+\tau_{iF}^f)w_F)y_i.$
2. Consumption taxes / wedges.
   - $P = 1/Y = p_0 = C_0((1+\tau_1^c)p_1, ..., (1+\tau_N^c)p_N).$
3. Endogenous wedges (e.g., endogenous markups).

Framework: Generality

1. Input-specific wedges.
   - $TC_i(y_i) = \frac{1}{A_i} C_i((1+\tau_{i1})p_1, ..., (1+\tau_{iN})p_N, (1+\tau^f_{i1})w_1, ..., (1+\tau^f_{iF})w_F)y_i.$
2. Consumption taxes / wedges.
   - $P = 1/Y = p_0 = C_0((1+\tau^c_1)p_1, ..., (1+\tau^c_N)p_N).$
3. Endogenous wedges (e.g., endogenous markups).
   - $Y = \mathcal{Y}(\mathcal{A}, \mathcal{X}(\mathcal{A}, \mu)) \Rightarrow Y = \mathcal{Y}(\mathcal{A}(h), \mathcal{X}(\mathcal{A}(h), \mu(\mathcal{A}, h))).$

Framework: Generality

1. Input-specific wedges.
   - $TC_i(y_i) = \frac{1}{A_i} C_i((1+\tau_{i1})p_1, ..., (1+\tau_{iN})p_N, (1+\tau^f_{i1})w_1, ..., (1+\tau^f_{iF})w_F)y_i$.
2. Consumption taxes / wedges.
   - $P = 1/Y = p_0 = C_0((1+\tau^c_1)p_1, ..., (1+\tau^c_N)p_N)$.
3. Endogenous wedges (e.g., endogenous markups).
   - $Y = \mathcal{Y}(\mathcal{A}, \mathcal{X}(\mathcal{A}, \mu)) \Rightarrow Y = \mathcal{Y}(\mathcal{A}(h), \mathcal{X}(\mathcal{A}(h), \mu(\mathcal{A}, h)))$.
4. Decreasing returns to scale. Interpretation of profits?
5. Input-augmenting technologies.
6. Elastic factor supply.

Input–Output Notation

- Now, we must keep track of both revenue-based and cost-based direct exposures:

$$\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}, \qquad \tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{\sum_k p_k x_{ik}} = \mu_i \Omega_{ij}.$$

Is $\sum_j \tilde{\Omega}_{ij} \lesseqgtr 1$? $\sum_j \Omega_{ij} \lesseqgtr 1$?

Input–Output Notation

- Now, we must keep track of both revenue-based and cost-based direct exposures:

$$\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}, \qquad \tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{\sum_k p_k x_{ik}} = \mu_i \Omega_{ij}.$$

Is $\sum_j \tilde{\Omega}_{ij} \lesseqgtr 1$? $\sum_j \Omega_{ij} \lesseqgtr 1$?

- Likewise, we now have revenue-based and cost-based indirect exposures,

$$\Psi = (I - \Omega)^{-1}, \qquad \tilde{\Psi} = (I - \tilde{\Omega})^{-1}.$$

- Revenue-based and cost-based Domar weights given by first row of Ψ and $\tilde{\Psi}$,

$$\lambda = \Psi^{(0)}, \qquad \tilde{\lambda} = \tilde{\Psi}^{(0)}.$$

- Which are sales shares? What happens if the economy is efficient ($\mu_i = 1$ for all i)?

Given a productivity shock, which Domar weight matters?

- How do prices respond to shocks?

$$d\log p_i = d\log \mu_i - d\log A_i + \sum_j \tilde{\Omega}_{ij} d\log p_j + \sum_f \tilde{\Omega}_{if} d\log w_f.$$

- Stacking, and using $d\log \Lambda_f = d\log w_f + d\log L_f = d\log w_f$, we have

$$d\log p = d\log \mu - d\log A + \tilde{\Omega} d\log p + \tilde{\Omega}_{(f)} d\log \Lambda.$$

$$\Rightarrow d\log p = \tilde{\Psi} \left(d\log \mu - d\log A + \tilde{\Omega}_{(f)} d\log \Lambda \right).$$

- So for output,

$$d\log Y = -d\log p_0 = \tilde{\lambda}' (d\log A - d\log \mu) - \tilde{\Lambda}' d\log \Lambda.$$

Theorem

- The effect of productivity and markup shocks $d\log A$ and $d\log\mu$ on output is

$$d\log Y = \underbrace{\tilde{\lambda}' d\log A}_{\Delta \text{ Technological Efficiency}} \quad \underbrace{-\tilde{\lambda}' d\log\mu - \tilde{\Lambda}' d\log\Lambda}_{\Delta \text{ Allocative Efficiency}}.$$

- $d\log\Lambda$ is how factor income shares change with shock (can be observed ex post).

- We will characterize $d\log\Lambda$ in terms of economy's structure (ex ante results).

- Since these results hold fixed factor supplies, growth accounting:

$$d\log Y = d\log \mathrm{TFP} + \tilde{\Lambda}' d\log L.$$

  where

$$d\log \mathrm{TFP} = \tilde{\lambda}' d\log A - \tilde{\lambda}' d\log\mu - \tilde{\Lambda}' d\log\Lambda.$$

Theorem: Starting from efficiency

$$d\log Y = \underbrace{\tilde{\lambda}' d\log A}_{\Delta \text{ Technological Efficiency}} \underbrace{-\tilde{\lambda}' d\log \mu - \tilde{\Lambda}' d\log \Lambda}_{\Delta \text{ Allocative Efficiency}}.$$

- What happens when economy is initially efficient?.
- Technology shock $d\log A_i$:

$$\begin{aligned} d\log Y &= \lambda_i d\log A_i - \tilde{\Lambda}' \frac{d\log \tilde{\Lambda}}{d\log A_i} d\log A_i \\ &= \lambda_i d\log A_i - \sum_f \frac{d\tilde{\Lambda}_f}{d\log A_i} d\log A_i \\ &= \lambda_i d\log A_i + 0. \end{aligned}$$

- Useful property: $\sum_f \tilde{\Psi}_{if} = 1$ for all i. Thus, $\sum_f \tilde{\Lambda}_f = 1$.

Theorem: Starting from efficiency

$$d\log Y = \underbrace{\tilde{\lambda}' d\log A}_{\Delta \text{ Technological Efficiency}} \underbrace{-\tilde{\lambda}' d\log \mu - \tilde{\Lambda}' d\log \Lambda}_{\Delta \text{ Allocative Efficiency}}.$$

- What happens when economy is initially efficient?.
- Markup shock $d\log \mu_i$:

$$\begin{aligned} d\log Y &= -\lambda_i d\log \mu_i - \tilde{\Lambda}' \frac{d\log \Lambda}{d\log \mu_i} d\log \mu_i \\ &= -\lambda_i d\log \mu_i - \frac{d\left(\sum_f \Lambda_f\right)}{d\log \mu_i} d\log \mu_i \\ &= -\lambda_i d\log \mu_i - (-\lambda_i)\, d\log \mu_i \\ &= 0. \end{aligned}$$

Theorem: Effect of Productivity and Markup Shocks

- In general, for productivity shocks

$$\frac{d\log Y}{d\log A_k} = \underbrace{\tilde{\lambda}_k}_{\Delta \text{ Technology}} \quad \underbrace{-\sum_f \tilde{\Lambda}_f \frac{d\log \Lambda_f}{d\log A_k}}_{\Delta \text{ Allocative Efficiency}} .$$

- For markup shocks,

$$\frac{d\log Y}{d\log \mu_k} = \underbrace{-\tilde{\lambda}_k - \sum_f \tilde{\Lambda}_f \frac{d\log \Lambda_f}{d\log \mu_k}}_{\Delta \text{ Allocative Efficiency}} .$$

Example: Vertical supply chain

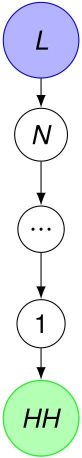


Figure: Vertical supply chain.

- Firm k has productivity A_k and markup μ_k.
- What is the labor share of income, Λ_L?

Example: Vertical supply chain

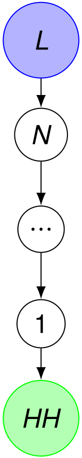


Figure: Vertical supply chain.

- Firm k has productivity A_k and markup μ_k.
- Labor share of income $\Lambda_L = 1/\prod_{i=1}^{N} \mu_i$.
- What is the effect of a markup shock on output?

Example: Vertical supply chain

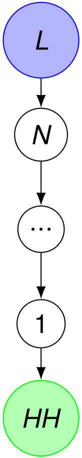


Figure: Vertical supply chain.

- Firm k has productivity A_k and markup μ_k.
- Labor share of income $\Lambda_L = 1/\prod_{i=1}^{N} \mu_i$.
- What is the effect of a markup shock on output?

$$d\log Y = -\tilde{\lambda}_i d\log\mu_i - \tilde{\Lambda}_L d\log\Lambda_L.$$

$$d\log\Lambda_L = -d\log\mu_i, \qquad \text{and} \qquad \tilde{\lambda}_i = \tilde{\Lambda}_L = 1.$$

Example: Vertical supply chain

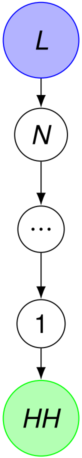


Figure: Vertical supply chain.

- Firm k has productivity A_k and markup μ_k.
- Labor share of income $\Lambda_L = 1/\prod_{i=1}^{N} \mu_i$.
- What is the effect of a markup shock on output?

$$d\log Y = -\tilde{\lambda}_i d\log \mu_i - \tilde{\Lambda}_L d\log \Lambda_L.$$

$$d\log \Lambda_L = -d\log \mu_i, \qquad \text{and} \qquad \tilde{\lambda}_i = \tilde{\Lambda}_L = 1.$$

- What is the effect of a productivity shock on output?

Example: Vertical supply chain

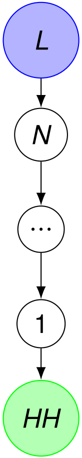


Figure: Vertical supply chain.

- Firm k has productivity A_k and markup μ_k.
- Labor share of income $\Lambda_L = 1/\prod_{i=1}^{N} \mu_i$.
- What is the effect of a markup shock on output?

$$d\log Y = -\tilde{\lambda}_i d\log \mu_i - \tilde{\Lambda}_L d\log \Lambda_L.$$

$$d\log \Lambda_L = -d\log \mu_i, \qquad \text{and} \qquad \tilde{\lambda}_i = \tilde{\Lambda}_L = 1.$$

- What is the effect of a productivity shock on output?
- What would Hulten's theorem have given us?

Example: Horizontal economy

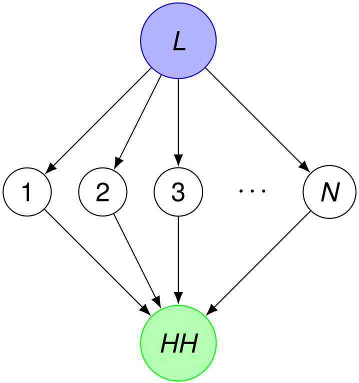


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

Example: Horizontal economy

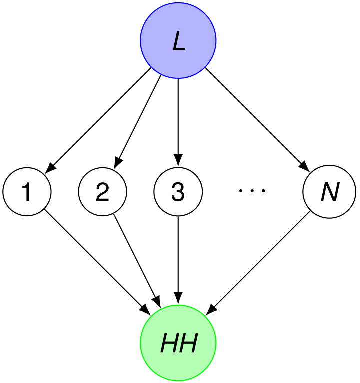


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d \log Y}{d \log \mu_i} = -\tilde{\lambda}_i - \tilde{\Lambda}_L \frac{d \log \Lambda_L}{d \log \mu_i}$$
$$= \lambda_i \theta_0 \left(\frac{\bar{\mu}}{\mu_i} - 1 \right),$$

where the sales-weighted harmonic average markup $\bar{\mu} = \left(\sum_{i=1}^{N} \lambda_i / \mu_i \right)^{-1} = 1/\Lambda_L$.

- What is the effect of a productivity shock on output?

Example: Horizontal economy

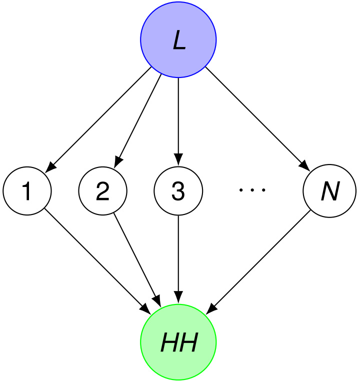


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d \log Y}{d \log \mu_i} = -\tilde{\lambda}_i - \tilde{\Lambda}_L \frac{d \log \Lambda_L}{d \log \mu_i}$$
$$= \lambda_i \theta_0 \left(\frac{\bar{\mu}}{\mu_i} - 1 \right),$$

where the sales-weighted harmonic average markup $\bar{\mu} = \left(\sum_{i=1}^{N} \lambda_i / \mu_i \right)^{-1} = 1/\Lambda_L$.

- What is the effect of a productivity shock on output?

$$\frac{d \log Y}{d \log A_i} = \tilde{\lambda}_i - \tilde{\Lambda}_L \frac{d \log \Lambda_L}{d \log A_i}$$
$$= \lambda_i - \lambda_i(\theta_0 - 1) \left(\frac{\bar{\mu}}{\mu_i} - 1 \right),$$

Example: Round-about economy

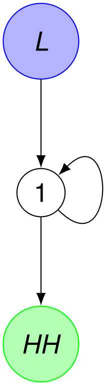


Figure: Round-about economy.

- Round-about firm uses L and x_{11} with elasticity θ_1.
- Effect of a markup shock:

$$\frac{d\log Y}{d\log \mu_1} = -\tilde{\lambda}_1 - \tilde{\Lambda}_L \frac{d\log \Lambda_L}{d\log \mu_1}$$
$$= \theta_1 \lambda_1 (\tilde{\lambda}_1 - 1)(1/\mu - 1).$$

- When is this positive? Negative?
- Effect of a productivity shock:

$$\frac{d\log Y}{d\log A_1} = \tilde{\lambda}_1 - \tilde{\Lambda}_L \frac{d\log \Lambda_L}{d\log A_1}$$
$$= \tilde{\lambda}_1 - (\theta_1 - 1)\lambda_1 (\tilde{\lambda}_1 - 1)(1/\mu - 1).$$

Beware of ad hoc productivity indices

- Many popular decompositions of changes in an ad-hoc aggregate productivity index.
- E.g., Baily, Hulten, and Campbell (1992) define a "TFP index"

$$\log A = \sum_i \lambda_i \log A_i.$$

$$d \log A = \sum_i \lambda_i d \log A_i + \sum_i d\lambda_i \log A_i.$$

- E.g., Olley and Pakes (1992) define "industry productivity" for an industry with N firms,

$$A = \bar{A} + \sum_{i=1}^{N} (A_i - \bar{A}) \left(\lambda_i - \frac{1}{N} \right),$$

where $\bar{A} = (1/N) \sum_i A_i$.

- These indices do not generally coincide with TFP, and the decompositions can detect efficiency gains from reallocation where there are none.

Alternative useful decomposition

- Another useful decomposition from Petrin and Levinsohn (2012):

$$d\log Y = \sum_i \lambda_i d\log A_i + \sum_i \lambda_i (1 - \mu_i^{-1})(d\log y_i - d\log A_i).$$

  Uses quantities rather than factor income shares.

- Careful with interpretation:
  - May be tempting to conclude that you always want to reallocate quantities toward high-markup producers.
  - The second term is not due to reallocations alone.

Caution: Interdependencies

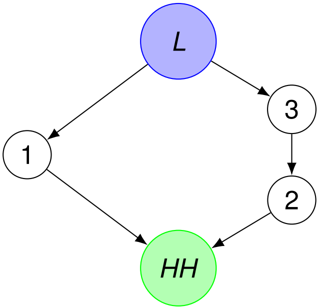


Figure: Interdepencies.

- In horizontal economy, high-wedge $=$ more distorted.
- This is not always the case.
- 1, 3 use labor directly. Suppose $\mu_1 > \mu_3 = 1$.
- Should we move labor to 1?
- Consider subsidy to labor for 1. We get:

$$\frac{d \log Y}{d \log \tau_1} = -\theta_0 \lambda_1 (1 - \lambda_1) \bar{\mu} \left[\frac{\mu_2 - \mu_1}{\mu_1 \mu_2} \right].$$

- Depends on $\mu_2 \lesseqgtr \mu_1$. Echoes McKenzie (1951) warning.

Caution: Spurious reallocation effects in vertical supply chain

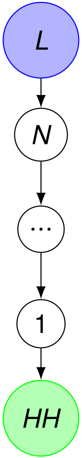
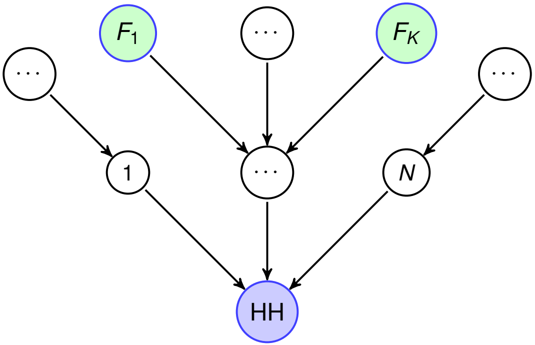


Figure: Vertical supply chain.

- Suppose we have productivity shock to firm N.
- $d\log Y = d\log A_N$, entire effect due to technology.
- Petrin and Levinsohn (2012) decomposition:

$$
\begin{aligned}
d\log Y &= \lambda_N d\log A_N + \sum_{i=1}^{N-1} \lambda_i (1-\mu_i^{-1}) d\log y_i \\
&= [\lambda_N + (\lambda_{N-1} - \lambda_N) + (\lambda_{N-2} - \lambda_{N-1}) + \ldots] d\log A_N \\
&= \lambda_N d\log A_N + (\lambda_1 - \lambda_N) d\log A_N \\
&= d\log A_N.
\end{aligned}
$$

Acyclic economies: No reallocation effects!



- Unique feasible allocation, hence efficient and no reallocation effects.
- Alternative decompositions find spurious changes in “allocative efficiency.”

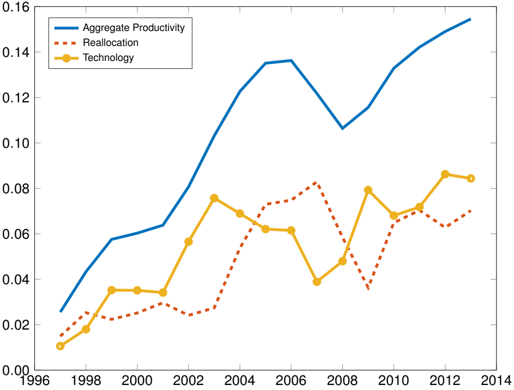
Growth accounting: Quantitative application to US markups

- How important are changes in allocative efficiency for US growth?
- Idea: Suppose markups are only distortions.
  - Take BEA input–output data from 1997–2015.
  - Assign Compustat firms to each industry.
  - Measure firms' sales shares and markups.
  - Use changes in factor income shares (labor and capital) in the data.
- Distortion-adjusted Solow residual

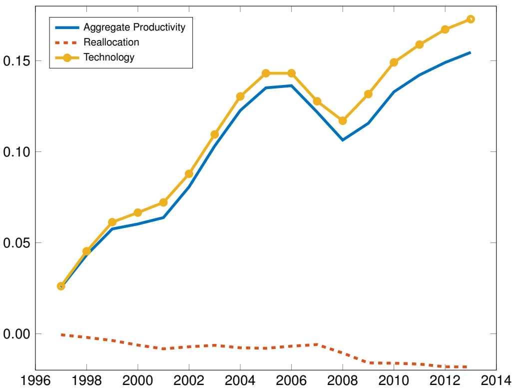
$$\Delta \log A_t \approx \Delta \log Y_t - \tilde{\Lambda}'_{t-1} \Delta \log \Lambda_t .$$

- Change in allocative efficiency, $-\tilde{\lambda}' \Delta \log \mu - \tilde{\Lambda}'_{t-1} \Delta \log \Lambda_t$.
- Residual is change in technology.

Growth accounting: Quantitative application to US markups



(a) Firm-level data.



(b) Industry-level data.

- Δ allocative efficiency accounts for half of growth.
- Within-industry reallocations to distorted producers (important to disaggregate!).

Table of Contents

Changes in factor shares

- How can we solve for changes in factor shares $d\log\Lambda$ from shocks?

- Define covariance operator,

$$Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(K)},\Psi_{(L)}\right)=\sum_i\tilde{\Omega}_{ji}\tilde{\Psi}_{iK}\Psi_{iL}-\left(\sum_i\tilde{\Omega}_{ji}\tilde{\Psi}_{iK}\right)\left(\sum_i\tilde{\Omega}_{ji}\Psi_{iL}\right).$$

- For productivity shocks, changes in factor shares solve the system,

$$\frac{d\log\Lambda_f}{d\log A_k}=\frac{1}{\Lambda_f}\sum_j\frac{\lambda_j}{\mu_j}(\theta_j-1)Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}-\sum_g\tilde{\Psi}_{(g)}\frac{d\log\Lambda_g}{d\log A_k},\Psi_{(f)}\right).$$

- Likewise, for markup shocks,

$$\frac{d\log\Lambda_f}{d\log\mu_k}=-\frac{1}{\Lambda_f}\sum_j\frac{\lambda_j}{\mu_j}(\theta_j-1)Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}+\sum_g\tilde{\Psi}_{(g)}\frac{d\log\Lambda_g}{d\log A_k},\Psi_{(f)}\right)-\lambda_k\frac{\Psi_{kf}}{\Lambda_f}.$$

Changes in factor shares: Single factor

- Let's consider a case with a single factor (labor):

- Productivity shocks:

$$\frac{d\log\Lambda_L}{d\log A_k} = \frac{1}{\Lambda_L}\sum_j \frac{\lambda_j}{\mu_j}(\theta_j - 1)Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right).$$

- Dependence on $Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}/\Lambda_L\right)$:
    - $\tilde{\Psi}_{(k)}$ measures how much price of i decreases when A_i increases.
    - $\Psi_{(L)}/\Lambda_L$ measures how high markups are on i's supply chain relative to overall economy.
    - If $Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}/\Lambda_L\right) < 0$, substitution by j lowers labor share.
    - Effects summed up over all producers weighted by their cost shares λ_j/μ_j.

- Why does this equal zero in an efficient economy?

Changes in factor shares: Single factor

- Let's consider a case with a single factor (labor):

- Markup shocks:

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}\sum_j \frac{\lambda_j}{\mu_j}(\theta_j - 1)Cov_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- Substitution effects analogous to productivity shocks.

- Now we also have direct effect on wage.

- How does this expression simplify when the economy is initially efficient?

Example: Horizontal economy

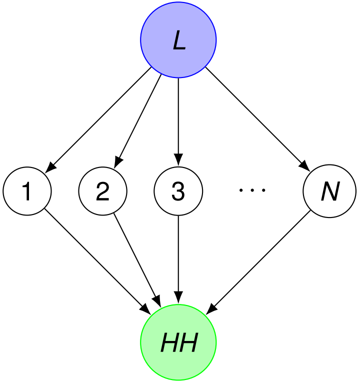


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$?

Example: Horizontal economy

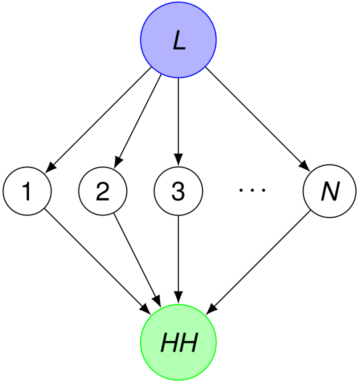


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.

Example: Horizontal economy

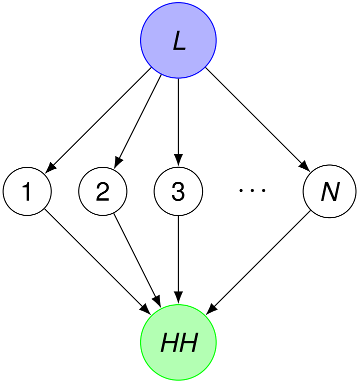


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.
- What is $\tilde{\Psi}_{ik}$?

Example: Horizontal economy

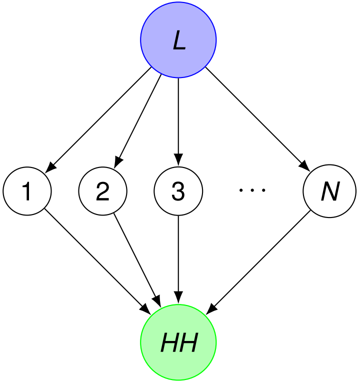


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.
- What is $\tilde{\Psi}_{ik}$? $1\{i = k\}$.

Example: Horizontal economy

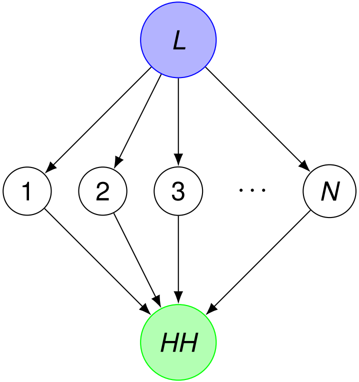


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.
- What is $\tilde{\Psi}_{ik}$? $1\{i = k\}$.
- What is Ψ_{iL}?

Example: Horizontal economy

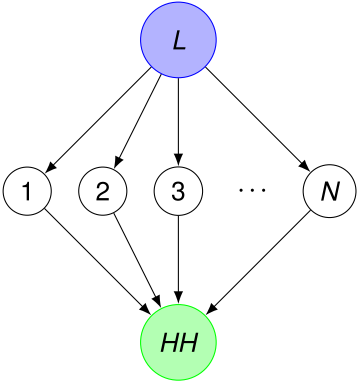


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.
- What is $\tilde{\Psi}_{ik}$? $1\{i = k\}$.
- What is Ψ_{iL}? $1/\mu_i$.

Example: Horizontal economy

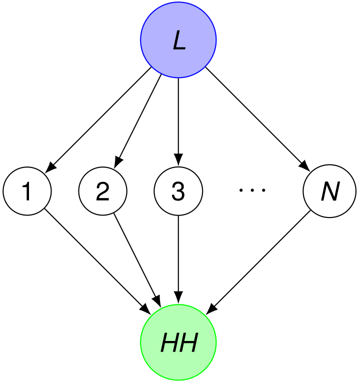


Figure: Horizontal economy.

- What is the effect of a markup shock on output?

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\frac{1}{\Lambda_L}(\theta_0 - 1)Cov_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(L)}\right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L}.$$

- What is $\tilde{\Omega}_{0i}$? λ_i.
- What is $\tilde{\Psi}_{ik}$? $1\{i = k\}$.
- What is Ψ_{iL}? $1/\mu_i$.

$$\frac{d\log\Lambda_L}{d\log\mu_k} = -\bar{\mu}(\theta_0 - 1)\left(\frac{\lambda_k}{\mu_k} - \lambda_k \sum_i \frac{\lambda_i}{\mu_i}\right) - \lambda_k \frac{\bar{\mu}}{\mu_k}$$

$$= -\theta_0 \lambda_k \left(\frac{\bar{\mu}}{\mu_k} - 1\right) - \lambda_k.$$

Table of Contents

Losses from markup distortions

- What are the gains of moving to the efficient frontier?

$$\text{Distance to frontier} = \frac{\log Y(\mathcal{A}, \mathcal{X}(\mathcal{A}, 1))}{\log Y(\mathcal{A}, \mathcal{X}(\mathcal{A}, \mu))}.$$

- Recall: Effect of reallocations at frontier is zero.

- At inefficient point, effect of changing quantities $\Delta \log y_i$ to a first order is

$$\Delta \log Y \approx \sum_i \lambda_i (1 - \mu_i^{-1})(\Delta \log y_i).$$

- Thus, distance to frontier given by summing over Harberger triangles:

$$\text{Distance to frontier} \approx \frac{1}{2} \sum_i \lambda_i (1 - \mu_i^{-1}) \Delta \log y_i \approx -\frac{1}{2} \sum_i \lambda_i (\Delta \log \mu_i)(\Delta \log y_i).$$

Losses from markup distortions

$$\text{Distance to frontier} \approx -\frac{1}{2}\sum_i \lambda_i (\Delta \log \mu_i)(\Delta \log y_i).$$

- We can solve for $\Delta \log y_i \approx \sum_k \frac{\partial \log y_i}{\partial \log \mu_k} \Delta \log \mu_k$ given network structure.
- For economy with single factor,

$$\text{Distance to frontier} \approx \frac{1}{2}\sum_i \lambda_i \theta_i Var_{\Omega^{(i)}}\left(\sum_k \Psi_{(k)} \Delta \log \mu_k\right).$$

Losses from markup distortions

$$\text{Distance to frontier} \approx -\frac{1}{2}\sum_i \lambda_i (\Delta \log \mu_i)(\Delta \log y_i).$$

- We can solve for $\Delta \log y_i \approx \sum_k \frac{\partial \log y_i}{\partial \log \mu_k} \Delta \log \mu_k$ given network structure.
- For economy with single factor,

$$\text{Distance to frontier} \approx \frac{1}{2}\sum_i \lambda_i \theta_i \, Var_{\Omega^{(i)}}\left(\sum_k \Psi_{(k)} \Delta \log \mu_k\right).$$

- For horizontal economy,

$$\text{Distance to frontier} \approx \frac{1}{2}\theta_0 \, Var_\lambda(\Delta \log \mu_i).$$

- Recall Hsieh and Klenow (2009): $\frac{\theta_0}{2} Var(\log \text{TFPR}_i)$.

Distance to frontier: Quantitative application to US economy

- BEA input–output data for 2015.
- Assign Compustat firms to each industry, and measure their sales and markups.
- Structural parameters:
  - Elasticities of substitution in consumption: 0.9.
  - Elasticity of substitution between value-added and intermediates: 0.5.
  - Elasticity of substitution across intermediates: 0.01.
  - Elasticity of substitution between labor and capital: 1.
  - Elasticity of substitution across firms outputs within industry: 8.

Distance to frontier: Quantitative application to US economy

Markup measure	Distance to the frontier User Cost (UC)	 Accounting (AP)	 Production Function (PF)
2015	13%	11%	25%
1997	3%	5%	23%

- Economy became more productive (growth accounting) but distance to frontier also increased.
- Cost of distortions: Contrast with 0.1% estimate of Harberger (1954) triangles.
- Tobin: *"It takes a heap of Harberger triangles to fill an Okun gap."*

Distance to frontier: Quantitative application to US economy

	Benchmark	CD + CES	$\xi = 4$	Cobb-Douglas	No I–O	Sectoral
UC	13%	14%	8%	3%	5%	0.7%
AP	11%	12%	6%	3%	5%	1%
PF	25%	29%	14%	10%	14%	4%

- Elasticities matter.
- Input-output structure matters.
- Illustrates importance of disaggregation.

Table of Contents

Misallocation in the New Keynesian Model

- Per-period utility:

$$u_t(Y_t, L_t) = \frac{Y_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi}, \qquad \text{where} \qquad Y_t = \left(\int_0^1 Y_t(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

subject to budget constraint,

$$P_t Y_t = W_t L_t + T_t, \qquad \text{where} \qquad P_t = \left(\int_0^1 P_t(i)^{1-\varepsilon} di \right)^{\frac{1}{1-\varepsilon}}.$$

- Production:

$$Y_t(i) = L_t(i).$$

- Prices including optimal subsidy τ and wedges from price rigidities, $\mu_i(t)$:

$$P_t(i) = \mu_t(i) \frac{\varepsilon}{\varepsilon - 1} \tau W_t.$$

Optimal subsidy $\tau = 1 - \varepsilon^{-1}$ sets $\frac{P_t Y_t}{W_t L_t} = 1$.

Misallocation in the New Keynesian Model

- Household optimality:

$$Y_t^{1-\sigma} = \frac{P_t Y_t}{W_t L_t} L_t^{1+\varphi}.$$

- Differentiating per-period utility to a second-order in shocks,

$$\begin{aligned}\Delta u_t = {} & Y_t^{1-\sigma} (\Delta \log Y_t) - L_t^{1+\varphi} (\Delta \log L_t) \\ & + \frac{1}{2} (1-\sigma) Y_t^{1-\sigma} (\Delta \log Y_t)^2 - \frac{1}{2} (1+\varphi) L_t^{1+\varphi} (\Delta \log L_t)^2 + h.o.t.\end{aligned}$$

- Change in output due to change in inputs and misallocation:

$$\Delta \log Y_t = \Delta \log L_t - \frac{\varepsilon}{2} Var (\log \mu_t (i)) + h.o.t.$$

- Plugging in and evaluating at optimal point,

$$\Delta u_t = -\frac{1}{2} Y_t^{1-\sigma} \left[\varepsilon \, Var (\log \mu_t (i)) + (\sigma + \varphi) (\Delta \log Y_t)^2 \right] + h.o.t.$$

Misallocation in the New Keynesian Model

$$\Delta u_t = -\frac{1}{2} Y_t^{1-\sigma} \left[\underbrace{\varepsilon \, Var\left(\log \mu_t\left(i\right)\right)}_{\substack{\text{Cost of inflation:} \\ \text{Misallocation due to price dispersion}}} + \underbrace{\left(\sigma + \varphi\right)\left(\Delta \log Y_t\right)^2}_{\substack{\text{Cost of output gap:} \\ \text{Moving off labor-leisure condition}}} \right] + h.o.t.$$

- $Var(\log \mu_t(i))$ depends on inflation due to price rigidity.
- Simple e.g.: Suppose wage jumps permanently by $\Delta \log W_t$. With Calvo friction δ,

$$Var(\log \mu_t(i)) \approx \delta(1-\delta)(\Delta \log W_t)^2 = \frac{1-\delta}{\delta}(\Delta \log \pi_t)^2.$$

  More generally, Woodford (2003, ch. 6) shows

$$\sum_{t=0}^{\infty} \beta^t Var(\log \mu_t(i)) = \frac{1-\delta}{\delta(1-\beta(1-\delta))} \sum_{t=0}^{\infty} \beta^t (d \log \pi_t)^2 .$$

Misallocation in the New Keynesian Model

- Costs of inflation in the canonical NK model are misallocation costs.
- Due to inefficient price/markup dispersion, which arises from price-setting frictions.
- Nakamura, Steinsson, Sun, and Villar (2018) on "The Elusive Costs of Inflation":
  - *In standard New Keynesian models [...] the consumption-equivalent welfare loss of moving from 0% inflation to 12% inflation is roughly 10%."*
  - Conclusions from measuring (a proxy for) price dispersion during the 1970s Great Inflation: *"There is thus no evidence that prices deviated more from their optimal level during the Great Inflation period when inflation was running at higher than 10% a year than during the more recent period when inflation has been close to 2% a year. We conclude from this that the main costs of inflation in the New Keynesian model are completely elusive in the data."*

Does inflation lead to inefficient price dispersion and misallocation?

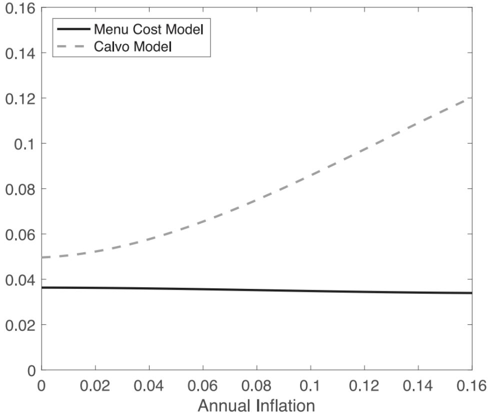


FIGURE V

Inefficient Price Dispersion

(a) Price dispersion vs. inflation.

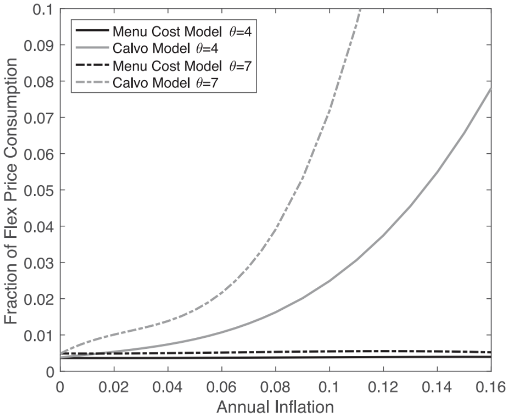


FIGURE II

Welfare Loss

(b) Welfare loss vs. inflation.

Source: Nakamura, Steinsson, Sun, and Villar (2018).

Does inflation lead to inefficient price dispersion and misallocation?

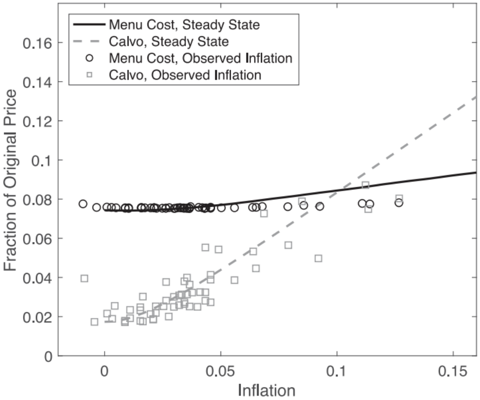


FIGURE VIII

Mean Absolute Size of Price Changes in Sticky-Price Models

(a) Absolute size of price changes: Model.

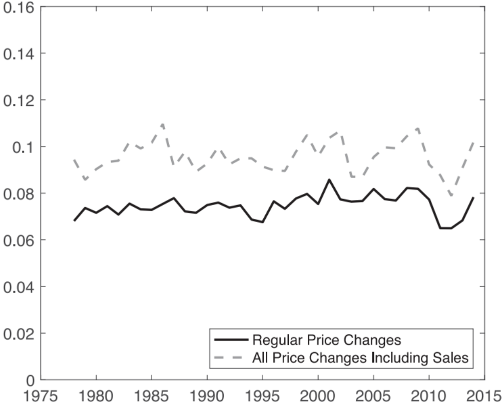


FIGURE IX

Absolute Size of Price Changes in U.S. Data

(b) Absolute size of price changes: Data.

Source: Nakamura, Steinsson, Sun, and Villar (2018).

Misallocation, revisited with data from Argentina's hyperinflation

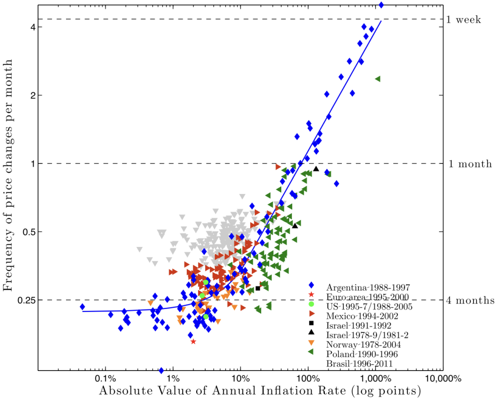


FIGURE VI

The Frequency of Price Changes (λ) and Expected Inflation: International Evidence

(a) Frequency of price changes.

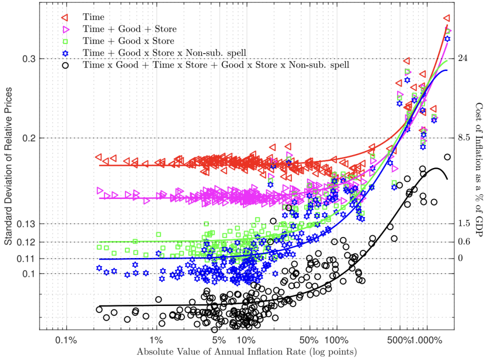


FIGURE X

Cross-Sectional Standard Deviation of Prices and Costs of Price Dispersion versus Inflation

(b) Price dispersion.

Source: Alvarez, Beraja, Gonzalez-Rozada and Neumeyer (2019).

Price dispersion: Micro to macro

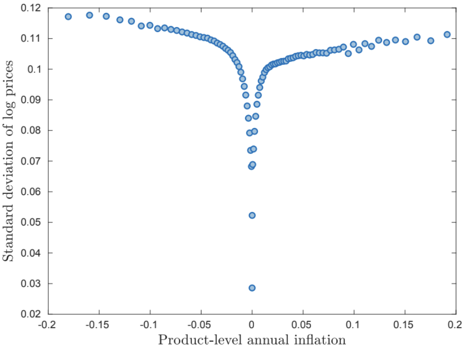


Figure 1: Product-level price dispersion and inflation, raw data

Figure: Sara-Zaror (2024).

- Nakamura et al. (2018) find little relationship between agg. inflation and price dispersion.
- At micro level, steep relationship between inflation and price dispersion!
- Aggregation over non-linear relationship.
- Alvarez et al. (2019) model with idiosyncratic productivity shocks and menu costs.