Productivity and Misallocation in General Equilibrium

David Baqaeel Emmanuel Farhi

UCLA Harvard

January 23, 2020
Aggregation Theorems for Efficient Economies

- Solow (1957) for economies with aggregate production functions:
  \[ d \log Y = d \log TFP + \sum_f \Lambda_f d \log L_f. \]

- Hulten (1978) for disaggregated economies with HA+IO:
  \[ d \log TFP = \sum_k \lambda_k d \log A_k, \quad \text{where} \quad \lambda_k = \frac{sales_k}{GDP}. \]

- Structural foundation for Domar aggregation, not definition.

- Measurement (growth accounting); predictions (counterfactuals).

- Powerful irrelevance result: disaggregated details (IO network, factors, returns to scale, elasticities, wealth distribution and mpcs); initial level of aggregation.
What We Do

- Extend these results to inefficient economies and other shocks.
- General reduced-form, non-parametric formula.
- Mapping from micro to macro using a *general* GE model:
  - micro wedges;
  - micro elasticites of substitution;
  - returns to scale;
  - factor market reallocation;
  - network linkages.

- Wide range of applications in different contexts: sources of TFP growth, impact of misallocation, macro impact of micro shocks, monetary policy with nominal rigidities, etc.

- Some selected numbers:
  - 50% of TFP growth 1997-2014 from improved allocative efficiency.
  - 20% rise in TFP from eliminating markups.
Related Literature

- **Efficient Network Production Economies:**

- **Inefficient Network Production Economies:**

- **Misallocation**
Related Literature

Agenda

General Result with Ex-Post Sufficient Statistics
  Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
  Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Agenda

General Result with Ex-Post Sufficient Statistics
   Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
   Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
General Framework

- Final demand as maximizer of homothetic aggregator:

\[ Y = D(c_1, \ldots, c_N), \]

with \( c_k \) final consumption of good \( k \).

- Budget constraint:

\[ \sum_{k} (1 + \tau^c_k) p_k c_k = \sum_{f} w_f F_f + \sum_{k} \pi_k + \tau, \]

with \( p_k \) prices, \( \pi_k \) profits, \( \tau^c_k \) consumption wedges, \( w_f \) wages, \( F_f \) factors, \( \tau \) lump-sum rebate.
General Framework

- Good $k$ produced with constant-returns cost function:

$$\frac{y_k}{A_k} C_k ( (1 + \tau_{k1}) p_1, \ldots, (1 + \tau_{kN}) p_N, (1 + \tau_{k1}^f) w_1, \ldots, (1 + \tau_{kF}^f) w_F ),$$

with $y_k$ total output, $A_k$ Hicks-neutral productivity shock, $\tau_{kl}$ input-specific wedge, $\tau_{ki}^f$ factor-specific wedge.

- Markup $\mu_k$ over marginal cost.

- Equilibrium: all markets clear.
Generality

- Captures factor augmenting productivity shocks with relabeling.
- Captures demand shocks as mix of productivity shocks.
- Captures decreasing returns with fixed quasi-factors.
- Captures increasing returns with fixed bad quasi-factors.
- Captures a form of entry/exit with choke prices.
- Can capture “technical” adjustment costs and capacity utilization.
- Can be applied to final demand within period, or intertemporally.
Notation and Accounting Convention

- Represent all wedges as markups with relabeling.

- Assume that in data, expenditures by $i$ on $j$ and revenues of $i$ recorded *gross* of wedges and markups.

- If not, for ex. with implicit wedges (e.g. credit constraints), re-write expenditures gross of these wedges.
Revenue-Based vs. Cost-Based

Definition

Ω and ˜Ω are $N \times N$ input-output matrices with $ij$th element:

$$
\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}, \quad \tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{\sum_k p_k x_{ik} + \sum_f w_f F_{if}}.
$$

Ψ and ˜Ψ are $N \times N$ Leontief inverse matrices:

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

$b$ is $N \times 1$ consumption-shares vector with $i$th element:

$$
b_i = \frac{p_i c_i}{\sum_j p_j c_j}.
$$

λ and ˜λ are $N \times 1$ Domar weights:

$$
\lambda = b' \Psi, \quad \tilde{\lambda} = b' \tilde{\Psi}.
$$
Revenue-Based vs. Cost-Based

Cost-based definitions capture correct notion of exposure:

- \( \tilde{\Omega}_{ij} \) is direct exposure of \( i \) to \( j \).
- \( \tilde{\Psi}_{ij} \) is direct and indirect exposure of \( i \) to \( j \).
- \( \tilde{\lambda}_k \) is direct and indirect exposure of household to \( k \).
Macro Impact of Micro Shocks

- \( \mathcal{Y}(A, X) \): output \( Y \) given productivities \( A \) and shares \( X_{ij} = x_{ij}/y_j \).

- Change in equilibrium in response to shocks:

\[
\text{d log } Y = \frac{\partial \log \mathcal{Y}}{\partial \log A} \text{d log } A + \frac{\partial \log \mathcal{Y}}{\partial X} \text{dX}.
\]

- \( \Delta \text{Technology} \) \quad \Delta \text{Allocative Efficiency}

- For efficient economies, macro envelope implies Hulten:

\[
\text{d log } Y = \lambda' \text{d log } A + 0.
\]

- \( \Delta \text{Technology} \) \quad \Delta \text{Allocative Efficiency}

- Inefficient economies: no macro envelope, only micro envelopes.
Macro Impact of Micro Productivity Shocks

Theorem

\[
\frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log A_k}.
\]

\(\Delta Technology\) \hspace{1cm} \(\Delta Allocative Efficiency\)

- Yields Hulten’s theorem for efficient economies:

\[
\tilde{\lambda}_k = \lambda_k \quad \text{and} \quad - \sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log A_k} = 0.
\]

- See later for structural formula for \(- \sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log A_k}\).
Macro Impact of Micro Markup Shocks

Theorem

$$\frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - \sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log \mu_k}.$$

$\Delta$Allocative Efficiency

- Also applies to shocks to other wedges.
- Can be applied to endogenous wedges via chain rule.
- See later for structural formula for $-\sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log \mu_k}$.
Ex. Simple Vertical Economy

- Example of multiple marginalization taken from Baqee (2016):
  \[ \tilde{\lambda}_k = 1 \neq \lambda_k = \prod_{i=1}^{k-1} \mu_i^{-1} \text{ and } \Lambda_L = \prod_{i=1}^{N} \mu_i^{-1} \neq 1. \]

- Productivity shocks:
  \[ \frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log A_k} = 1 \]

- Markups/wedges shocks:
  \[ \frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log \mu_k} = 0 \]
Ex. Simple Horizontal Economy

- $\tilde{\lambda}_k = \lambda_k$ and $\Lambda_L = \sum_j \lambda_j \mu_j^{-1} \neq 1$.

- Productivity shocks:
  
  \[
  \frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log A_k} = \lambda_k - (\theta_0 - 1) \left( \frac{\mu_k^{-1}}{\sum_j \lambda_j \mu_j^{-1}} - 1 \right) \lambda_k.
  \]

- Markup/wedge shocks:
  
  \[
  \frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log \mu_k} = \theta_0 \left( \frac{\mu_k^{-1}}{\sum_j \lambda_j \mu_j^{-1}} - 1 \right) \lambda_k.
  \]
Growth Accounting

- Change in aggregate TFP as new “distorted” Solow residual:
  \[ d \log TFP = d \log Y - \tilde{\Lambda}' d \log L. \]

- Decomposition of changes in aggregate TFP:
  \[ d \log TFP = \tilde{\lambda}' d \log A - \tilde{\lambda}' d \log \mu - \tilde{\Lambda}' d \log \Lambda. \]
  
  \text{pure technology} \quad \text{allocative efficiency}

- Can perform decomposition without imposing \textit{any} parametric assumptions on production functions.

- Generalizes Hall (88,90) for disaggregated economies.
Alternative Decompositions: Statistical

- Popular decompositions: Baily et al. (92), Giriliches-Regev (95), Olley-Pakes (96), Foster et al. (01).

- Decompositions of change in ad-hoc aggregate TFP index.

- Not decompositions of change aggregate TFP.

- Ex. Baily et al. (92):

  \[ d \log \left( \sum_i \lambda_i A_i \right) = \sum_i \lambda_i d \log A_i + \sum_i A_i d \log \lambda_i, \]
Alternative Decompositions: Economic


- Ad-hoc decompositions of change in aggregate TFP.

- “Grouping of terms”, not GE counterfactuals.

- Ex. Jorgenson et al. (1987):

\[
\begin{align*}
\text{d log } TFP &= \sum_i \lambda_i \text{d log } A_i + \left( \text{d log } TFP - \sum_i \lambda_i \text{d log } A_i \right).
\end{align*}
\]
Alternative Decompositions: Misleading

- Detect reallocation effects when they unambiguously shouldn’t:
  - efficient economies;
  - economies without reallocation.

- See also Osotimehin (19).
Revealing Example of Acyclic Economies

Unique feasible allocation, hence efficient.

No reallocation effects, no changes in allocative efficiency.

Alternative decompositions fail.
Agenda

General Result with Ex-Post Sufficient Statistics
   Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
   Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Application: Markups in US

- Suppose markups are only distortions.
- Use annual IO tables from BEA from 1997-2015.
- Assign Compustat firms to industries.
- Use firm-level markups from three approaches: user cost, production function, and accounting profits.
- Aggregate-up from firm level.
(Harmonic) Average Markups: Between and Within

- With user-cost-approach markup data.
- Similar with other approaches for markups.
Sources of Growth

- With user-cost-approach markup data.
- Similar with other approaches for markups.
Sources of Growth: Industry Level Instead of Firm Level

- With user-cost-approach markup data.
- Similar with other approaches for markups.
- Illustrates importance of disaggregation.
General Result with Ex-Post Sufficient Statistics
  Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
  Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Parametric Model

- General nested-CES economy with wedges.
- Relabel network so that each node corresponds to one CES nest.
- Today: assume a single factor (see paper for multiple factors).

**Definition**

\[
\text{Cov}_{\tilde{\Omega}(j)} \left( \tilde{\Psi}_k, \Psi_L \right) = \sum_i \tilde{\Omega}_{ji} \tilde{\Psi}_{ik} \Psi_{iL} - \left( \sum_i \tilde{\Omega}_{ji} \tilde{\Psi}_{ik} \right) \left( \sum_i \tilde{\Omega}_{ji} \Psi_{iL} \right).
\]
Macro Impact of Micro Productivity Shocks: One Factor

Proposition

Suppose there is only one factor (with index L). Then

\[
\frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log A_k},
\]

\[
= \tilde{\lambda}_k - \sum_j (\theta_j - 1) \mu_j^{-1} \lambda_j \text{Cov}_{\tilde{\Omega}(j)} \left( \tilde{\Psi}_k, \frac{\Psi(L)}{\Lambda_L} \right).
\]

- Change in allocative efficiency opposite of change in labor share.
- Centrality measure mixing network and elasticities.
- Upstream and downstream distortions matter.
Explaining Covariance Operator

\[
\frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \sum_j (\theta_j - 1) \mu_j^{-1} \lambda_j \text{Cov}_{\tilde{\Omega}^{(j)}} \left( \tilde{\Psi}_k, \frac{\Psi(L)}{\Lambda_L} \right).
\]

- High \( \tilde{\Psi}_{ik} \): \( i \)'s highly exposed to \( k \).
- High \( \Psi_{iL}/\Lambda_L \): most of \( i \)'s revenues are ultimately paid to workers.
Change in technology and change in allocative efficiency:

\[ \frac{d \log Y}{d \log A_k} = \lambda_k - (\theta_0 - 1) \left( \frac{\mu_k^{-1}}{\sum_j \lambda_j \mu_j^{-1}} - 1 \right) \lambda_k. \]

Key: markup vs. average and elasticity minus one.
Macro Impact of Micro Markup Shocks: One Factor

Proposition

Suppose there is only one factor indexed by $L$. Then

$$
\frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - \frac{d \log \Lambda_L}{d \log \mu_k},
$$

which is equal to

$$
\frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - \left[ \sum_j (1 - \theta_j) \mu_j^{-1} \lambda_j \text{Cov}_{\tilde{\Omega}_j} \left( \tilde{\Psi}_k, \frac{\Psi(L)}{\Lambda_L} \right) - \lambda_k \frac{\Psi_{kL}}{\Lambda_L} \right].
$$

- First two terms like a negative productivity shock.
- Third term captures that increase in markups releases labor.
Change in allocative efficiency:

\[
\frac{d \log Y}{d \log \mu_k} = -\tilde{\lambda}_k - (1 - \theta_0)\lambda_k \left(\frac{\mu_k^{-1}}{\Lambda_L} - 1\right) + \frac{\lambda_k \mu_k^{-1}}{\Lambda_L},
\]

\[
= \theta_0 \left(\frac{\mu_k^{-1}}{\sum_j \lambda_j \mu_j^{-1}} - 1\right) \lambda_k.
\]

Key: markup vs. average and elasticity.
Macro Impact of Micro Productivity Shocks: Multiple Factors

The following linear system describes the elasticities of factor shares:

\[
\frac{d \log \Lambda}{d \log A_k} = \Gamma \frac{d \log \Lambda}{d \log A_k} + \delta^{(k)},
\]

with

\[
\Gamma_{F,L} = \sum_j (\theta_j - 1) \lambda_j \mu_j^{-1} \text{Cov}_{\tilde{\Omega}(j)} \left( \tilde{\Psi}(F), \frac{\Psi(L)}{\Lambda_L} \right),
\]

and

\[
\delta^{(k)}_F = \sum_j (\theta_j - 1) \lambda_j \mu_j^{-1} \text{Cov}_{\tilde{\Omega}(j)} \left( \tilde{\Psi}(k), \frac{\Psi(F)}{\Lambda_F} \right).
\]

Given the elasticities of factor shares, we have

\[
\frac{d \log Y}{d \log A_k} = \tilde{\lambda}_k - \sum_f \tilde{\Lambda}_f \frac{d \log \Lambda_f}{d \log A_k}.
\]

• Similar for markup/wedge shocks.
Ex. Multiple Factors
Measuring Distance to Frontier

- Distance to frontier focus of recent misallocation literature (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009):

\[ L = \log \left( \frac{Y(A,1)}{Y(A,\mu)} \right). \]

- Can be computed by cumulating changes in allocative efficiency along a path to the frontier using our measure:

\[
L = - \int_0^1 \frac{d \log Y(A,\hat{\mu}(t))}{d \log \mu} \frac{d \log \hat{\mu}(t)}{d t} dt \\
= - \frac{1}{2} \sum_i \frac{d \log Y(A,\mu)}{d \log \mu_i} \log \mu_i + O(\| \log \mu \|^3),
\]

where \( \log \hat{\mu}_k(t) = \tau \log \mu_k \).
Distance to Frontier: Second-Order Approximations

- Sales-share weighted sum of Harberger triangles (ex post):
  \[ \mathcal{L} \approx - \sum_j \frac{1}{2} \lambda_j \Delta \log \mu_j \Delta \log y_j. \]

- Structural formula (ex ante)...ex. for one-factor (generalizes):
  \[ \mathcal{L} \approx \sum_j \frac{1}{2} \lambda_j \theta_j \operatorname{Var}_{\Omega(j)} \left( \sum_k \psi_{(k)} \Delta \log \mu_k \right). \]

- Generalizes Hsieh-Klenow formula: markups/wedges, elasticities, input-output network, and their joint distribution.
Comparison to Hsieh-Klenow

- Distance to frontier for horizontal economy:
  \[ \mathcal{L} \approx \frac{1}{2} \theta_0 \text{Var}_\lambda (\Delta \log \mu). \]

- Boils down to Hsieh-Klenow formula if \((A_i, \mu_i)\) lognormal:
  \[ \mathcal{L} \approx \frac{1}{2} \theta_0 \text{Var}(\Delta \log \mu). \]

- Correlation \(\lambda_i\) or \(A_i\) vs. \(\mu_i\) matters in general.

- Our formula captures it but Hsieh-Klenow’s doesn’t.
Alternative Decompositions with Different Objectives

- **Our decomposition:**

\[
d \log Y = \frac{\partial \log Y}{\partial \log A} \, d \log A + \frac{\partial \log Y}{\partial X} \, d X.
\]

\[\Delta \text{Technology} \quad \Delta \text{Allocative Efficiency}\]

- **Debreu-Farrell:**

\[
d \log Y = d \log Y^* + (d \log Y - d \log Y^*).
\]

\[\Delta \text{Technology} \quad \Delta \text{Allocative Efficiency}\]

- **Osotimehin:**

\[
d \log Y = \left[ \frac{\partial \log Y}{\partial \log A} + \frac{\partial \log Y}{\partial X} \frac{\partial X}{\partial \log A} \right] d \log A + \frac{\partial \log Y}{\partial X} \frac{\partial X}{\partial \log \mu} d \log \mu.
\]

\[\Delta \text{Technology} \quad \Delta \text{Allocative Efficiency}\]

- **Alternative decompositions can be computed with our structural formulas, but require more knowledge of the structure of the economy (elasticities of substitution).**
General Result with Ex-Post Sufficient Statistics
   Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
   Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Application: Gains from Eliminating Markups

- Calibrate parametric model.

- Use IO table from BEA from 2015.

- Benchmark elasticities of substitution: across industries in consumption 0.9; between value-added and intermediates 0.5; across intermediates in production 0.01; between labor and capital 1; within industries 8.
Gains from Eliminating Markups in US

<table>
<thead>
<tr>
<th></th>
<th>User Cost (UC)</th>
<th>Accounting (AP)</th>
<th>Production Function (PF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>13%</td>
<td>11%</td>
<td>25%</td>
</tr>
<tr>
<td>1997</td>
<td>3%</td>
<td>5%</td>
<td>23%</td>
</tr>
</tbody>
</table>

- Measures show big increase between 1997 and 2014.
- Contrast with 0.1% estimate of Harberger (1954) triangles.

“It takes a heap of Harberger triangles to fill an Okun gap.” — Tobin
**Gains from Eliminating Markups: Robustness**

<table>
<thead>
<tr>
<th></th>
<th>Benchmark</th>
<th>CD + CES</th>
<th>$\xi = 4$</th>
<th>Cobb-Douglas</th>
<th>No IO</th>
<th>Sectoral</th>
</tr>
</thead>
<tbody>
<tr>
<td>UC</td>
<td>13%</td>
<td>14%</td>
<td>8%</td>
<td>3%</td>
<td>5%</td>
<td>0.7%</td>
</tr>
<tr>
<td>AP</td>
<td>11%</td>
<td>12%</td>
<td>6%</td>
<td>3%</td>
<td>5%</td>
<td>1%</td>
</tr>
<tr>
<td>PF</td>
<td>25%</td>
<td>29%</td>
<td>14%</td>
<td>10%</td>
<td>14%</td>
<td>4%</td>
</tr>
</tbody>
</table>

- Elasticities matter.
- Input-output structure matters.
- Illustrates importance of disaggregation.
General Result with Ex-Post Sufficient Statistics
  Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
  Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Other Applications (see paper)

- Macro impact of micro shocks.
- Macro volatility from micro shocks.
- Sticky prices, monetary policy, and productivity.
Theoretical Extensions (see paper)

- Endogenous markups/wedges.
- Elastic Factors.
- Entry.
- Nonlinearities.
- Heterogenous households.
General Result with Ex-Post Sufficient Statistics
   Application: Growth Accounting

General Result with Ex-Ante Sufficient Statistics
   Application: Gains from Eliminating Markups in US

Extensions (see paper)

Conclusion
Conclusion

- Ex-post aggregation theorems for economies with frictions.
- Ex-ante aggregation theorems for economies with frictions.
- Wide range of applications in different contexts.
- Work in progress: structural models of frictions (IO, financing constraints, search and matching, nominal rigidities, etc.), fixed costs, entry and exit, dynamics, non-homotheticities, endogenous innovation, other models of network formation, etc.
- Part of a broader research agenda on disaggregated heterogeneous production vs. aggregate production function.


Autor, D., D. Dorn, L. Katz, C. Patterson, and J. Van Reenen (2017). The fall of the labor share and the rise of superstar firms.


