Macro as Explicitly Aggregated Micro

David Baqaee   Emmanuel Farhi

The Aggregation Problem

- A program dating back to the very beginnings of economics.
- Derailed in 70s:
  - “negative” theoretical results (anything goes);
  - lack of disaggregated data.
- Recent explosion, building on pioneers of 80s and 90s.
- Increasing scope, scale, and availability of disaggregated data.
- Need flexible aggregation theories to harness deluge of data.

Macro as Explicitly Aggregated Micro

- (HA) Distribution of heterogeneous agents.
  consumption and factor supplies

- (IO) Input-output network of heterogenous producers.
  production from factors and intermediates

- Objective: develop a general approach to consistent theory and measurement of propagation and aggregation with HA+IO.

- Research agenda David Baqaee.

Talk Based on Series of Papers with David Baqaee

- Inefficient economies.

  "*Productivity and Misallocation in General Equilibrium*"

- Nonlinearities in efficient economies.

  "*Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem*"

- Economies with heterogeneous agents.

  "*Macroeconomics with Heterogeneous Agents and Input-Output Networks*"

- Intermediate-level aggregation.

  "*A Short Note on Aggregating Productivity*"

- Open economies.

  "*Networks, Barriers, and Trade*"

- Micro-foundations of aggregate production functions and the Cambridge-Cambridge Capital controversy.

  "*The Microeconomic Foundations of Aggregate Production Functions*"

- Ongoing work on increasing returns, entry/exit, dynamics, industry structure, growth, etc.

First-Order Aggregation Theorems for Efficient Economies

- Solow (1957) with aggregate production function:

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

- Hulten (1978) 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.

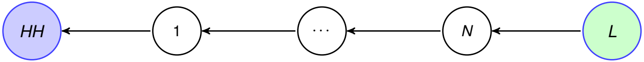
- Measurement (growth accounting); predictions (counterfactuals).

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

Limits of Hulten's Theorem and Need for New Theories

- Fails in inefficient economies.
- Fails at higher-orders of approx. relevant for nonlinearities.
- Disaggregated details and initial aggregation level matter.
- Need new theories for inefficient and nonlinear aggregation.

Why Sales Shares? Ex. Simple Vertical Economy



- k transforms $k-1$ linearly, productivity A_k.
- Zero value added for $k \neq N$ but sales share $\lambda_k = 1$.
- Productivity shock to k:

$$\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k} = \lambda_k = 1.$$

- More complex economies with substitution and reallocation?

Pure Technology Effects and Reallocation Effects

- Productivities A_i and allocation matrix $\mathscr{X}_{ij} = x_{ij}/y_j$ give allocation.

  exogenous factor supplies (generalizes)

- Aggregate output function $\mathscr{Y}(A, \mathscr{X})$.

  via homothetic final demand aggregator or at constant prices

- Change in equilibrium aggregate output in response to shocks:

$$d \log Y = \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \log A} d \log A}_{\text{pure technology}} + \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d \mathscr{X}}_{\text{reallocation}}.$$

Hulten's Theorem as an Envelope Theorem

$$d\log Y = \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A}_{\text{pure technology}} + \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X}}_{\text{reallocation}},$$

with

$$\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A = \lambda' d\log A \quad \text{and} \quad \frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X} = 0.$$

- Pure technology effects: Domar-weights sufficient stat.
- Zero reallocation effects despite reallocations (envelope).

Distortions

- Capture arbitrary distortions with saturating wedges.
- Can represent wedges as markups μ_i via network relabelling.
- Equilibrium aggregate output $Y(A,\mu) = \mathscr{Y}(A, \mathscr{X}(A,\mu))$.

Two Reasons for Failure of Hulten with Distortions

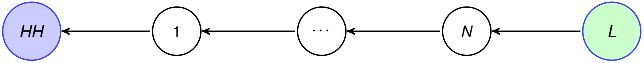
$$d\log Y = \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A}_{\text{pure technology}} + \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X}}_{\text{reallocation}},$$

with

$$\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A \neq \lambda' d\log A \quad \text{and} \quad \frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X} \neq 0.$$

- Two legs of Hulten's theorem break.
- Two key illustrating examples.

Ex. Simple Vertical Economy



- k transforms $k-1$ linearly, productivity A_k, markup/wedge μ_k.
- Productivity shock to k.

Ex. Simple Vertical Economy

$$\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k} = \underbrace{\frac{\mathrm{d}\log \mathscr{Y}}{\mathrm{d}\log A_k}}_{\text{pure technology}} + \underbrace{\frac{\mathrm{d}\log \mathscr{Y}}{\mathrm{d}\,\mathscr{X}}\frac{\mathrm{d}\,\mathscr{X}}{\mathrm{d}\log A_k}}_{\text{reallocation}}.$$

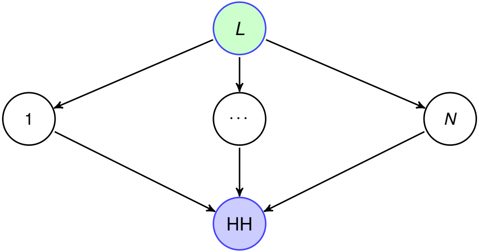
- Pure changes in technology:

$$\frac{\mathrm{d}\log \mathscr{Y}}{\mathrm{d}\log A_k} = \tilde{\lambda}_k = 1 \neq \lambda_k = \prod_{i=1}^{k-1} \mu_i^{-1}.$$

- Reallocation effects:

$$\frac{\mathrm{d}\log \mathscr{Y}}{\mathrm{d}\,\mathscr{X}}\frac{\mathrm{d}\,\mathscr{X}}{\mathrm{d}\log A_k} = 0.$$

Ex. Simple Horizontal Economy



- k produces from labor, productivity A_k, markup/wedge μ_k.
- CES consumption aggregator with elasticity θ_0.
- Allocation matrix $\mathscr{X}(A,\mu) = \left(\frac{L_1(A,\mu)}{L}, \cdots, \frac{L_N(A,\mu)}{L}\right)$.
- Productivity shock to k.

Ex. Simple Horizontal Economy

$$\frac{d\log Y}{d\log A_k} = \underbrace{\frac{d\log \mathscr{Y}}{d\log A_k}}_{\text{pure technology}} + \underbrace{\frac{d\log \mathscr{Y}}{d\mathscr{X}}\frac{d\mathscr{X}}{d\log A_k}}_{\text{reallocation}}.$$

- Pure changes in technology:

$$\frac{d\log \mathscr{Y}}{d\log A_k} = \lambda_k.$$

- Reallocation effects as changes in allocative efficiency:

$$\frac{d\log \mathscr{Y}}{d\mathscr{X}}\frac{d\mathscr{X}}{d\log A_k} = -\frac{d\log \Lambda_L}{d\log A_k} = -(\theta_0 - 1)\left(\frac{\mu_k^{-1}}{\sum_j \lambda_j \mu_j^{-1}} - 1\right)\lambda_k.$$

HA+IO concepts

- Revenue-based and cost-based input-output matrix:

$$\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}, \quad \tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{\sum_{j'} p_{j'} x_{ij'}}.$$

- Revenue-based and cost-based Leontief inverse matrix:

$$\Psi = (I - \Omega)^{-1} = I + \Omega + \Omega^2 + \ldots, \quad \tilde{\Psi} = (I - \tilde{\Omega})^{-1} = I + \tilde{\Omega} + \tilde{\Omega}^2 + \ldots$$

- Revenue-based and cost-based Domar weights:

$$\lambda = b' \Psi, \quad \tilde{\lambda} = b' \tilde{\Psi}.$$

- Revenue-based and cost-based factor shares Λ and $\tilde{\Lambda}$.

Aggregation with Distortions

$$d\log Y = \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A}_{\text{pure technology}} + \underbrace{\frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X}}_{\text{reallocation}}$$

with

$$\frac{\partial \log \mathscr{Y}}{\partial \log A} d\log A = \tilde{\lambda}' d\log A \quad \text{and} \quad \frac{\partial \log \mathscr{Y}}{\partial \mathscr{X}} d\mathscr{X} = -\tilde{\lambda}' d\log \mu - \tilde{\Lambda}' d\log \Lambda.$$

- Pure technology: cost-based Domar weights sufficient stat.
- Reallocation: changes in factor shares sufficient stat.
- Disaggregated details and intial aggregation level matter.

Propagation Equations

- Propagation equations:

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

- Can also derive equations for prices and quantities.
- IO network and elasticities sufficient stat.
- Can be extended and applied with endogenous wedges.

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:

$$d \log TFP = \underbrace{\tilde{\lambda}' d \log A}_{\text{pure technology}} \underbrace{- \tilde{\lambda}' d \log \mu - \tilde{\Lambda}' d \log \Lambda}_{\text{allocative efficiency}}.$$

- Can perform decomposition without imposing *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):

$$\mathrm{d}\log\left(\sum_i \lambda_i A_i\right) = \sum_i \lambda_i \,\mathrm{d}\log A_i + \sum_i A_i \,\mathrm{d}\log \lambda_i,$$

Alternative Decompositions: Economic

- Popular decompositions: Jorgenson et al. (1987), Basu-Fernald (2002), and Petrin-Levinsohn (2012).

- Ad-hoc decompositions of change in aggregate TFP.

- “Grouping of terms”, not GE couterfactuals.

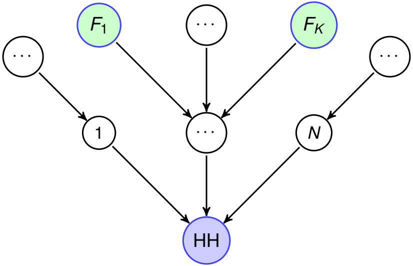
- Ex. Jorgenson et al. (1987):

$$\mathrm{d}\log TFP = \sum_i \lambda_i \,\mathrm{d}\log A_i + \left(\mathrm{d}\log TFP - \sum_i \lambda_i \,\mathrm{d}\log A_i \right).$$

Alternative Decompositions: Misleading

- Detect reallocation effects when they unambigously shouldn't:
  - efficient economies;
  - economies without reallocation.
- See also Osotimehin (19).

Revealing Example of Acyclic Economies

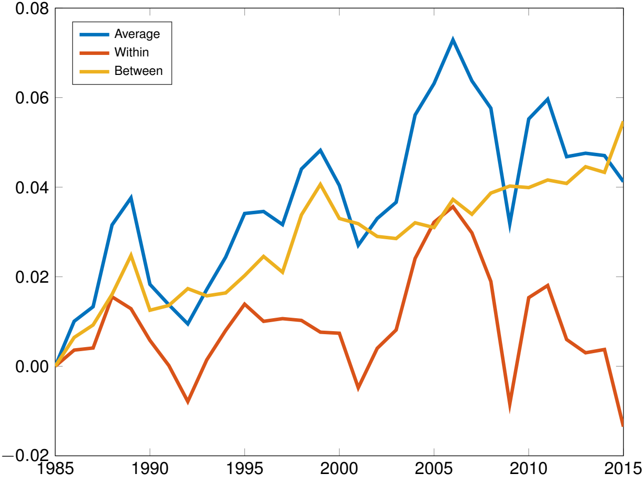


- Unique feasible allocation, hence efficient.
- No reallocation effects, no changes in allocative efficiency.
- Alternative decompositions fail.

Application: Markups in US

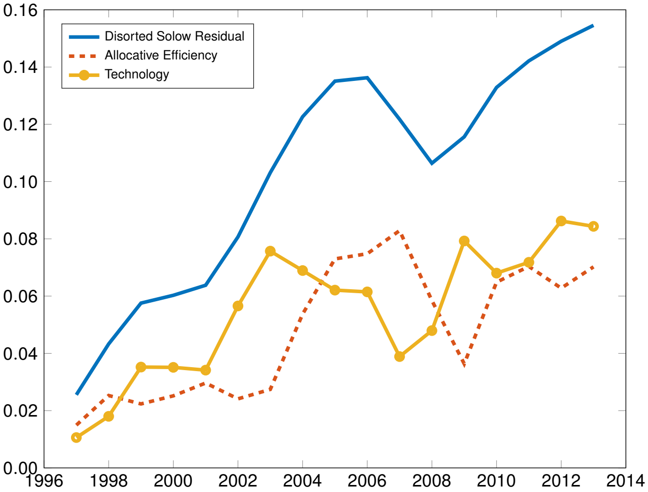
- Suppose markups are only distortions.
- Use annual IO tables from BEA from 1997-2015.
- Assign Compustat firms to industries.
- Use firm-level markups from three approaches: user cost, production function, and accounting profits.
- Aggregate-up from firm level.

(Harmonic) Average Markups: Between and Within



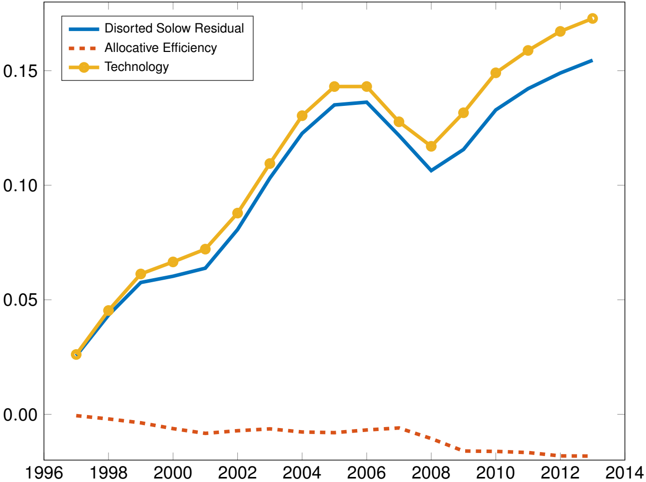
- With user-cost-approach markup data.
- Similar with other approaches for markups.

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.

Distance to the Frontier

- Different notion of change in allocative efficiency.
- Second-order approx. of distance to frontier for small wedges.
- Sum of Harberger triangles in GE:

  ∼Harberger (64) but aggregate TFP (not welfare) without implausible transfers

$$\mathscr{L} \approx -\sum_j \frac{1}{2} \lambda_j \Delta \log \mu_j \Delta \log y_j.$$

- Structural formula:

$$\mathscr{L} \approx \sum_j \frac{1}{2} \lambda_j \theta_j Var_{\Omega^{(j)}}\left(\sum_k \Psi_{(k)} \Delta \log \mu_k\right).$$

- Role of elasticities, IO network, distribution of wedges.

Application: Gains from Eliminating Markups in US

- 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

	User Cost (UC)	Accounting (AP)	Production Function (PF)
2015	20%	17%	24%
1997	3%	5%	17%

- 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

	Benchmark	CD + CES	$CES = 4$	No IO	Sectoral
UC	20%	21%	10 %	7 %	0.7%
AP	17%	18%	9 %	7 %	1%
PF	24 %	27 %	13%	13%	3%

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

Nonlinearities

$$d^2 \log Y = d \log A' \frac{\partial^2 \log \mathscr{Y}}{\partial \log A^2} d \log A + d \log A' \frac{\partial^2 \log \mathscr{Y}}{\partial \log A d \mathscr{X}} d \mathscr{X},$$

$$= d\lambda' d \log A \quad \text{with } d\lambda \text{ from propagation equations.}$$

- Second-order approx. (with $\Delta\lambda$ from propagation equations):

$$\Delta \log Y \approx \lambda' \Delta \log A + \frac{1}{2} \Delta\lambda' \Delta \log A.$$

- Changes in sales shares ex-post sufficient stat.
- IO network and elasticities ex-ante sufficient stat.
- Disaggregated details and initial aggregation level matter.

Back to Horizontal Economy

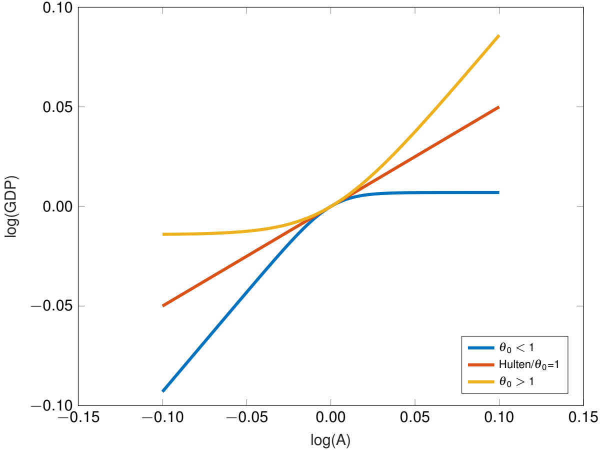
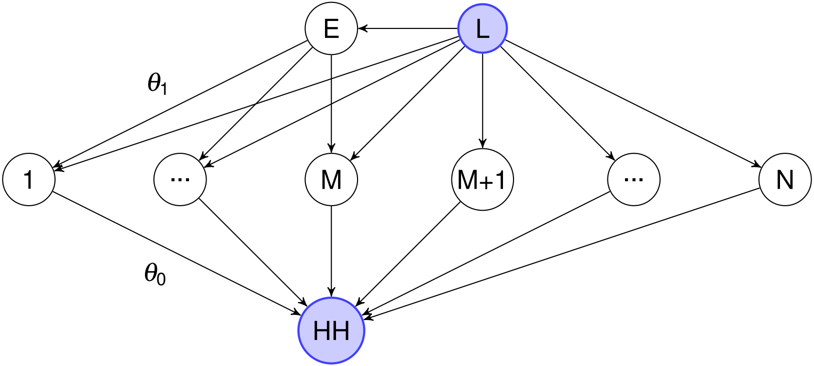


Figure: Aggregate output as a function of shocks to one of the producers.

- Cobb-Douglas knife-edge case with no nonlinearities.
- Implications for firm- vs. sector-level shocks.

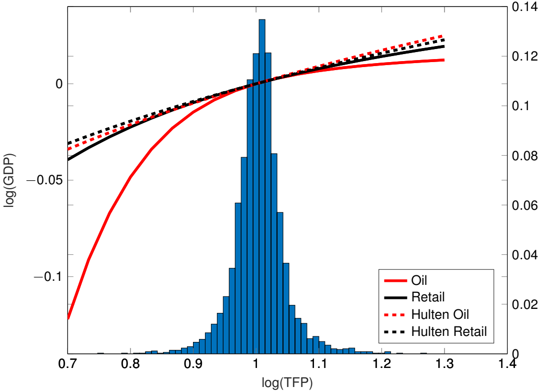
"Universal" Input Example

One factor, full reallocation, two elasticities $\theta_1 \ll \theta_0$.



$$\frac{\mathrm{d}^2 \log C}{\mathrm{d} \log A_E^2} = \frac{d\lambda_E}{d \log A_E} = (\theta_0 - 1)\lambda_E \left(\frac{N}{M} - 1 \right) \lambda_E + (\theta_1 - 1)\lambda_E \left(1 - \frac{N}{M} \lambda_E \right).$$

Oil v. Retail



- Simulated model, no markups, sectoral-level shocks.
- Intuition: low micro-elasticity of substitution, universal input.
- Large asymmetric effects of oil shocks (Hamilton, 2003).

Reduced-form Impact of Oil Shocks

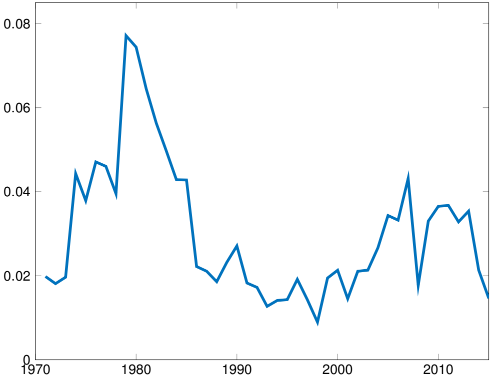


Figure: Global expenditures on crude oil as a fraction of world GDP.

- First-order effect: $1.8\% \times -13\% \approx -0.2\%$.
- Second-order effect: $\frac{1}{2}(1.8\% + 7.6\%) \times -13\% \approx -0.6\%$.

Baumol's Cost Disease and US TFP Growth

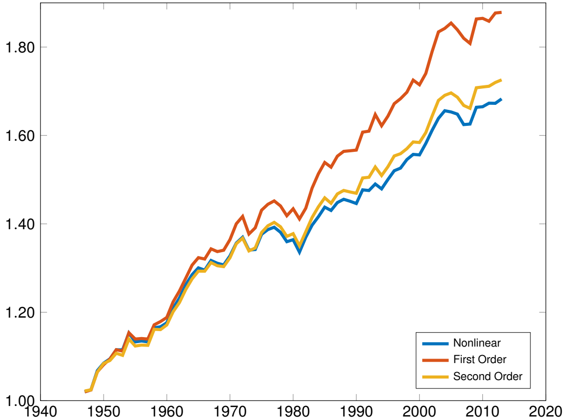


Figure: Cumulative change in *TFP*: nonlinear (actual), first-order approximation, and second-order approximation.

- Baumol's cost disease reduced aggregate TFP growth by 20 pp.

Application: Gains from Trade in Trade Models

(σ, ζ, θ)	No IO $(1, 1, 1)$	$(1, 1, 1)$	$(1, 0.5, 0.6)$	$(0.9, 0.5, 0.2)$
FRA	9.8%	18.5%	24.7%	30.2%
JPN	2.4%	5.2%	5.5%	5.7%
MEX	11.5%	16.2%	21.3%	44.5%
USA	4.5%	9.1%	10.3%	13.0%

- Accounting for IO network and cross-industry elasticities drastically magnifies gains from trade.
- Shortcuts with no IO fine qualitatively but not quantitatively.
- Illustrates importance of disaggregation.

Conclusion

- Macro as explicitly aggregated micro increasingly possible.
- More and better disaggregated data.
- New theories of propagation and aggregation with HA+IO.
- Explosion of work, much more needed (theory, data, empirics).