How to integrate micro-evidence into macro models?

Question: “Can micro data be used to identify behavioral elasticities and build models of individual behavior that can be usefully aggregated into macro models?”

- I took as: How to integrate micro evidence into macro modeling?
- Timely: Age of Big Data and methods for identification/prediction!
- Conceptually important to move the literature forward:
  - High-quality micro-data evidence is often informative about PE
  - But important macro-policy questions often depend on GE

My position: (Proper) small models, (properly) disciplined by data

- “More data” is better. “More/bigger model” is not always better....
Old Keynesian macro-econometric models (pre Lucas-critique)

First generation(s) DSGE models (before the Great Recession)

Computable general equilibrium models in international trade
  - Tim Kehoe: “Ex-post evaluations of the performance of applied GE models are essential if policy makers are to have confidence in the results produced by this sort of model.”

Covid-macro models with a quantitative focus
  - Robust lessons are qualitative. Did we need the big models?
Lessons from (mis)adventures of big macro models

Narrow lesson: Add more ingredients (RE, financial frictions, HANK...)

Broader lesson: Macro phenomena are enormously complicated

- “Model of everything” is tempting but can be misleading
- Different questions/mechanisms require different models

- Appropriate small models can also capture quantitative GE forces

Suppose we have quantitative ambitions (bring in micro evidence): Is a 17-equation GE model necessarily better than 4-equation?
Why small? Reality is messier than what we know to model

- Rich heterogeneity/interactions, dynamics, informational/behavioral frictions (rational expectations shortcut is useful but...)
- For a similar fit, might as well opt for simplicity/transparency
Recent recessions/crises seem like Tolstoy’s unhappy families:

- **Small, less ambitious** models designed for **specific mechanisms**
Insights from statistical learning: Model selection

- Statistical learning has a similar problem: Bias-variance trade-off
- Complex models tend to have smaller bias but more sampling variance
- Key idea: **Penalize complexity** to reduce variance (regularization)
- Key idea: **Model-selection techniques** (cross-validation)

We need: **Principles of model-building** that **penalizes complexity**
Model building: General principles

- Start with a **mechanism/question** (do NOT start with a framework)
- Start with a **small model** and add ingredients (NOT the opposite)
- Add **central ingredients** that matter for mechanism qualitatively
  - Measure these ingredients & estimate related elasticities carefully
- For more **tangential ingredients**, need to use judgement:
  - **Noise**: Is there strong micro evidence for the ingredient?
  - **Relevance**: How much does it matter for the mechanism?
  - **Shortcuts**: Can I capture essence with simpler ingredient (as-if)?
Example: Macro effects of stock market wealth effect

Chodorow-Reich, Nenov, Simsek (AER, 2021): Regional variation to identify **local GE effects of stock wealth**. Purpose of model:

- Roughly quantify the **aggregate GE effects**
- Roughly quantify the **implied household-level MPCs**

Central ingredients:

- Regions with heterogeneous stock wealth (try to measure well)
- Nontradables and tradables (qualitatively different response)
- Nominal rigidity
- Monetary policy at the aggregate level (affects the macro response)
- A Keynesian multiplier (with bounding argument, aggregate > local)

Result: Aggregate response (w/ passive policy) > Local **NT** response
Judgement call ingredients are open to debate

Endogenous stock prices, but with no aggregate risk \textbf{(as-if)}

- Shouldn’t we explain the Equity Premium Puzzle? No

Minimal household-heterogeneity, mostly to hit multiplier \textbf{(as-if)}

- Didn’t we learn from GFC that MPC-heterogeneity is important?
- Doesn’t it matter for the mechanism: Stocks are held by the wealthy?
- Upon closer look, this heterogeneity isn’t central in our context:
  - Most wealth inequality is \textit{within-counties} rather than \textit{across}
  - Simulations: County weighted-average MPC $\approx$ Aggregate MPC

Infinite horizons, though effectively two horizons \textbf{(tractability)}

- Don’t we need dynamics? Perhaps, but costly and not our focus

- Shouldn’t we add: Financial frictions, housing sector, fiscal policy, luxury goods, geographic spillovers...? Not necessarily
Example: Permanent income inequality and savings

Straub (2019):
- Consumption elasticity to **permanent income** is 0.7 (textbook = 1).
- Model with **non-homothetic** preferences. \( \uparrow \) Income inequality \( \implies \uparrow \) Aggregate wealth (& inequality), \( \downarrow \) Equilibrium interest rate
- Quantitative model to assess the magnitudes of macro effects

Central ingredients:
- Income elasticity that rises with age (the novel non-homotheticity)
- Bequests (a known, competing source of non-homotheticity)
- Heterogeneous skill distribution, and inheritance of parental skill
- Sensible aggregate wealth dynamics (absent non-homotheticity)
- Asset supply and its interest-elasticity (to gauge \( \Delta r \))

Judgement-call ingredients: Build on “canonical” life-cycle model
- **Big package** with idiosyncratic/persistent shocks, taxes/transfers...
- Pro: Can tap into existing knowledge. Con: Not strictly necessary
A bigger model, but carefully measured small/central core

Table 4: Calibrated parameters of the baseline non-homothetic life cycle model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth, death, skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Number of permanent types</td>
<td>3</td>
<td>see text</td>
</tr>
<tr>
<td>( {\mu, \nu} )</td>
<td>Population shares by type</td>
<td>{0.9, 0.09, 0.01}</td>
<td>see text</td>
</tr>
<tr>
<td>( {\delta} )</td>
<td>Mortality rates by age</td>
<td></td>
<td>CDC, 2011</td>
</tr>
<tr>
<td>Production</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>Capital share</td>
<td>0.37</td>
<td>NIPA, 2014</td>
</tr>
<tr>
<td>( {\gamma} )</td>
<td>Labor income shares</td>
<td>{0.65, 0.24, 0.11}</td>
<td>Piketty and Saez (2003); updated to 2014</td>
</tr>
<tr>
<td>( \delta )</td>
<td>Depreciation</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>( A )</td>
<td>Total factor productivity</td>
<td>0.63</td>
<td>match ( k / y = 3.05 ) (NIPA, 2014)</td>
</tr>
<tr>
<td>Government</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Federal debt held by the public / GDP</td>
<td>0.73</td>
<td>NIPA, 2014</td>
</tr>
<tr>
<td>( \tau^b )</td>
<td>Bequest tax</td>
<td>0.10</td>
<td>see text</td>
</tr>
<tr>
<td>( y )</td>
<td>Income floor</td>
<td>0.30W</td>
<td>literature</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>Income tax progressivity</td>
<td>0.18</td>
<td>PSID, 2013</td>
</tr>
<tr>
<td>( \tau^{intax} )</td>
<td>Average income tax</td>
<td>0.30</td>
<td>NIPA, see text</td>
</tr>
<tr>
<td>( \tau^{cap} )</td>
<td>Capital tax</td>
<td>0.40</td>
<td>NIPA, see text</td>
</tr>
<tr>
<td>( G/Y )</td>
<td>Government spending / GDP</td>
<td>0.13</td>
<td>govt. budget constraint</td>
</tr>
<tr>
<td>Productivities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho )</td>
<td>Income shock persistence</td>
<td>0.90</td>
<td>PSID</td>
</tr>
<tr>
<td>( \sigma_\rho^2 )</td>
<td>Var. of innovations to persistent shock</td>
<td>0.028</td>
<td>PSID</td>
</tr>
<tr>
<td>( \sigma_\sigma^2 )</td>
<td>Var. of transitory income shocks</td>
<td>0.055</td>
<td>PSID</td>
</tr>
<tr>
<td>( \sigma_\epsilon^2 )</td>
<td>Var. of measurement error in incomes</td>
<td>0.02</td>
<td>literature; see text</td>
</tr>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>Discount factor</td>
<td>0.89</td>
<td>match interest rate ( r = 0.03 )</td>
</tr>
<tr>
<td>( \vartheta )</td>
<td>Elast. of intertemp. substitution, median age</td>
<td>2.5</td>
<td>literature</td>
</tr>
<tr>
<td>( z )</td>
<td>Scale term in utility function</td>
<td>0.30</td>
<td>30% of average income</td>
</tr>
<tr>
<td>( \sigma_\beta )</td>
<td>Ratio of elasticities ( \sigma_{\beta+1}/\sigma_k )</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>( \kappa )</td>
<td>Weight on bequest motive</td>
<td>15.84</td>
<td>match bequests / GDP = 0.05</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Intercept in bequest utility</td>
<td>1.72</td>
<td>30% share with beq. &lt; 6.25% avg. income</td>
</tr>
</tbody>
</table>

Micro evidence: 
- Piketty and Saez (2003); updated to 2014
- Chetty et al. (2014)

Elasticity from empirical section: 
- match \( \gamma = 0.059 \)
A bigger model, but with a thoughtful discussion

Justify the complexity:

- “Why do I allow for both sources of non-homothetic consumption-savings behavior, and not merely focus on non-homotheticity in bequests?...because bequest flows are typically...around 5% of GDP, limiting their quantitative role.”

Beware of hidden central ingredients:

- $F$ is Cobb-Douglas. Matters for asset supply elasticity and $\Delta r$
- Justified on tradition, not micro data. Hard to estimate.
- Ludwig’s solution:
  - Show also PE response for given $\Delta r$
  - Analyze also alternative setup with inelastic assets/Lucas tree
Solow (JEP, 2008): “My general preference is for **small, transparent, tailored models**, often partial equilibrium, **usually aimed at understanding some little piece of the (macro-)economic mechanism**. I would also be for **broadening the kinds of data that are eligible for use in estimation and testing**. One of the advantages of this alternative style of research is that it should be easier to accommodate relevant empirical regularities derived from behavioral economics as they become established.”