Quantitative Macroeconomics

Calibration and the use of Data in Macroeconomics

Tomás R. Martinez

UnB

Introduction

- A discussion on the use of data in macro models and empirical strategies in macro.
- Most of the discussion here is my own thoughts on the topic (that changes constantly over time) and mostly applied to het. agents models.
    - Some material came from the notes of Matthias Kredler.

References

- Nakamura and Steinsson (2018, JEP): Modern review of the "Identification in Macro".
- Canova (2007, Book), Fernández-Villaverde et al (2009): Great references for estimation in Macro. Chapter 7 of Canova has a very nice discussion on "calibration".
- Chodorow-Reich (2020, JEDC), Guren et al. (2021, NBER Macro). Good references for the cross-regional empirical methods. Check Chodorow-Reich class notes as well.

Model

- **What is a model?**
  - **Sargent**: define a model as a probability distribution over a sequence of outcomes, possibly indexed by a parameter vector;
  - **Wikipedia:** a theoretical construct representing economic processes by a set of variables and a set of logical and/or quantitative relationships between them.
  - **Statistics:** a data generating process.
- **Why use models?**
- We are usually interested in *counterfactuals*: i.e. what happens to Y when we change X?
  - **Applied Micro:** counterfactual comes from natural experiments + statistical assumptions.
  - **Quantitative Macro:** counterfactual comes from a set of structural assumptions using economic theory and the parameters chosen for the models.
- Must choose a set of parameters θ to make counterfactuals.

Choosing parameters

- How we identify parameter values often matters a great deal for results in quantitative work.
- A method that is (or tries to be) fully transparent and has a cookbook recipe: maximum-likelihood.
- Another approach is to use some type of moment-based estimation, i.e. GMM, minimum distance.
- Often authors refer to choosing parameters as **calibration**. What is the difference?

Likelihood based

- Systematically uses all moments (full information).
- Used for DSGE models $\Rightarrow$ few data points!
- Use Bayesian estimation to make feasible.
- But: Have to choose which data series to match $\Rightarrow$ Similar issue arises as for moment-matching – what to choose?

Calibration

- What calibration might mean (from Canova's book):
  - "...one wants to calibrate a model (in the sense of selecting reasonable parameters values) because there is no data to estimate its parameters."
  - "..one may prefer to calibrate (as opposed to estimated it) if the misspecification is so large that statistical estimation of its parameters will produce inconsistent and/or unreasonable estimates and formal statistical testing will lead to outright rejection."
  - "...some users interpret calibration as an econometric technique where the parameters are estimated using "economic", as opposed to "statistical", criteria."
- "the term calibration is used to indicate a particular collection of procedures designed to provide an answer to economic questions using "false" models."

Calibration

The essence of calibration by Kydland and Prescott (1991, 1996):

(i) Choose an economic question to be addressed.

(ii) Select a model design which bears some relevance to the question asked.

(iii) Choose functional forms for the primitives of the model and find a solution for the endogenous variables in terms of the exogenous ones and of the parameters.

(iv) Select parameters and convenient specifications for the exogenous processes and simulate paths for the endogenous variables.

(v) Evaluate the quality of the model by comparing its outcomes to a set of ”stylized facts” of the actual data.

(vi) Propose an answer to the question, characterize the uncertainty surrounding the answer and do policy analyses if required.

Calibration

- Still many questions:
  - What are the set of "stylized facts"?
  - What is the measure of distance used to compare the model with actual data?
- A lot of discretionary choices by the research.
- Here it lies important philosophical aspect of the methodology:
  - In a strict sense, all models are approximations to the DGP and, as such, false and unrealistic;
  - Once this point of view is accepted, it makes no sense to examine the validity of a model using standard statistical tools which assume it to be true, at least under the null.
  - This is what is implicitly assumed with GMM and ML.

Calibration

- The modern definition of calibration can be summarized by this set of tweets from Jon Steinsson.
- **Jon Steinsson**: "Calibration is just moment matching without standard errors."
- Std. errors are important, but parameter uncertainty is minor relative to model misspecification.
  - Sometimes is too computationally cost to calculate standard errors (especially in heterogeneous-agents economies).
  - Not everybody agrees with this point.
- Model evaluation by calibration gives rise to "portable statistics", i.e., statistics that are used over and over again to evaluate different models.

Calibration

- The moments are generated after a full solution of the model (as opposed to only one FOC or eq. equation).
  - In some sense, this type of moment matching is similar to ML since requires solving the full model.
  - Calibration gives the freedom to choose which moments to match, in ML the freedom is selecting the set of observables to use for estimation.
- Nowadays, we are moving from moments that are just averages/correlations/variances, to matching causal estimates:
  - e.g., marginal propensity to consume out of a transitory fiscal rebate;

Models as over-identifying Restrictions

- A good theory imposes restrictions on the data.
- These restrictions/predictions can be tested and falsified.
- **Example**: Have normal distributions of body height for n countries: $\{\hat{\mu}, \hat{\sigma}\}_{i=1}^{n}$
- The model $x_i \sim N(\mu_i, \sigma_i)$ with $2n$ parameters is not particularly interesting: no degrees of freedom!
- Better: Posit that μ_i is a function of some covariate at country level: $\mu_i = \alpha x_i$, removes $n-1$ parameters.
- Perfect theory: Few parameters, but can match all moments.

The name of the game (usually)

- Calibrate m free model parameters to m targets.
- Then show k additional (non-targeted) moments in model and data:“model validation”.
- Fit should be decent to have a good model.
- **Example:** Target levels of an economy in 2000, then see how model does for period 2000-2020.
- Alternative (not done much): Could choose m moments to minimize distance to $m + k$ moments.
  - Perfect model: Cannot reject over-identifying restrictions statistically $\Rightarrow$ almost impossible in practice in economics!
  - OK model: We feel the quantitative economic fit of the moments is satisfactory to use the model to do policy evaluation etc.

Calibration in Practice

(i) Determine as many parameters as possible directly from data ("outside model").

(ii) Take (uncontroversial) parameters from other studies.

(iii) Calibrate the remaining parameters to match moments (exactly or over-identified method of moments).

(iv) Validate your model using non-target moments.

Step 1: Parameters directly identified outside model

Examples:

- Processes the model takes as exogenous and that we can estimate from data:
  - Income process in heterogeneous-agents model.
  - TFP in business cycles models.
  - Exogenous process for exit in firm dynamics models.

- Variables that can be taken from institutional environment/prices:
  - tax rates, tax schedules, tariffs.
  - relative prices for goods pin down technology in simple settings: $y = An$.

Step 2: Take parameters from other studies

- Kydland and Prescott (1991): choosing parameters using information obtained from other studies imposes coherence among various branches of the profession.

- Prime example: CES-utility curvature parameter ($\gamma = 2$).

- But careful: Does the same Greek letter really mean the same in my model and the other? Often not!

- **Example**: Risk aversion
  - measured in lab experiment with small sums;
  - risk aversion coefficient of representative agent in RBC model.

- A good compromise is to test the robustness of your results in an interval of these parameters consistent with empirical estimates.
  - e.g.: $\gamma \in [0.5, 1, 2, 3]$.

Step 3: Moment-matching

- What is often done: Minimize percentage differences.

$$\min_{\theta} \sum_{i=1}^{n} W_i [\ln \mu_i(\theta; \psi) - \ln \hat{\mu}]^2$$

  where ψ is a vector of "deep parameters" chosen in the previous step and W_i some arbitrary weight.

  - resembles usual moment-based estimators like SMM;
  - in fact, if we use some statistical criteria to minimize the loss function (i.e., minimize MSE, choose W_i efficiently, etc), the two methods are exactly the same;
  - the difference is in the discretionary choice of the calibrator;
  - **Canova**: "...a calibrator may look like an econometrician who uses different loss functions in different parts of the model"
  - "...a calibrator may also look like as an inefficient GMM econometrician."
- **Important**: the parameters θ are *conditional* on ψ.

Detour: Indirect Inference

- Another approach similar to SMM is **indirect inference**.
  - Simulate model data;
  - Estimate an auxiliary model (e.g., a linear regression) in both the simulated and actual data;
  - Minimize the distance between the coefficients of the **auxiliary model**.
- Useful when the structural relationships are difficult to express as simple unconditional moments.
- **Examples:**
  - Guvenen and Smith (E2014, ECTA): Consumption-savings with uncertainty about income-process.
  - Search-friction models with wage dispersion (e.g., Lise (2013, ReStud)).
- How to choose the auxiliary model?

Step 3: Moment-matching

Practical Issues:

- Hard task computationally
  - If moments smooth in θ: Use gradient-based method.
  - Otherwise: Use more robust methods (simplex, genetic algorithms, etc).
  - If too slow/impossible: try around to understand how parameters change the moments generate by the model.
- **Key challenge**: Which moments to pick?
  - informative moments: Which statistics are especially affected by a certain parameter?
  - Often hardest, but most important: elasticity-type (curvature) parameters. (labor-supply elasticity, risk aversion,...)

Informative moments

- Before estimating: Draw comparative-statics graphs with your model.
- Which parameter affects which moment most? Use this to moment to pin down parameter.
- Can also choose by economic reasoning on model properties.
- Good papers have discussion on identification (usually no proof since no direct one-to-one match).

Elasticity/curvature moments

- Determines how much agents change behavior when prices/incentives change.
- Can use cross-sectional distribution to pin this down.
- Ríos-Rull: "Don't identify elasticity/slope moment by a level moment".
- **Causal Effects as Identified Moments**
  - Target a "causally" estimated by the applied micro literature;
  - Target an impulse response function.
  - The advantage of identified moments is that they can provide evidence on specific causal mechanisms of a model and may be relatively invariant to other model features.

Other useful tricks

- Re-parameterize your model to have an unconstrained optimization problem. Examples:
- $\sigma = \exp(\hat{\sigma})$ for $\sigma > 0$ or $\beta = logist(\hat{\beta})$ for $\beta \in (0,1) \Rightarrow \hat{\sigma}$ and $\hat{\beta}$ live on entire real line
- Include equilibrium conditions in the loss function.
  - Example: Net demand of assets must be zero in equilibrium.
  - Penalize (net demand)^2; put harsher penalties than on other moments.
  - Useful when equilibrium-finding loop is computationally costly.
- Smooth discrete choices by adding preference shocks
  - Choice probabilities instead of 0-1.
  - Makes moments smoother in parameters.

Step 4: Model validation

- There are no free parameters and no uncertainty is allowed, so how to validate the model?
- Typical approach: use non-target moments
- **Examples**:
  - Kaplan-Moll-Violante (HANK): calibrate share of hand-to-mouth, validate using moments of wealth distribution;
  - Midrigan-Xu: calibrate productivity process, validate using autocorrelations of investment and employment.
- Bonus (hard) approach: replicate empirical studies!
  - Berger-Herkenhoff-Mongey (2022, AER): GE model of firms' labor market power.
  - Replicate "natural experiments" from empirical papers.

Step 4: Model validation

- It is useful to assess whether the moments are able to identify the parameters.
- A simple test is to "perturb" a parameter and compute how "sensible" is the loss function to changes in parameters.
  - Perturb each parameter (one at a time) by 1% (or any small value);
  - Compute the % change of the loss function relative to the value evaluated at the "estimated parameters"
- if the model is well identified, the loss function should not be flat in the region around the vector of estimated parameters.
- You can even go further and compute the % change for the contribution of **each moment** to the total loss.

Building Empirical Evidence

- Having theoretically sounding models in macro are nice but we still need empirical evidence to back-up our results.

- Macroeconomists have traditionally used time-series to identify fiscal and monetary shocks (among other). Often relies on some type of identified VAR:
  - Structural restrictions, i.e., restrictions of response of variables to shocks;
  - Sign-restrictions;
  - Narrative approach, e.g.., wars, oil shocks, identified tax changes;

- **Last 20 years**: tons of development in the applied micro literature $\Rightarrow$ macro researchers should embrace the causal-revolution!

Data and Methods in Macro

Table 10: Econometric Methods and Data Types Over Time

	Methods		Data				
Year Year	Time series	Applied micro	Micro data	Time series	Cross section	Panel	Proprietary
1980	75	25	22	89	8	3	13
1990	62	38	28	70	14	16	32
2000	58	42	28	54	8	38	30
2006–10	46	54	41	42	13	45	41
2016–18	35	65	56	34	10	56	52

Note: The figures are the shares, expressed as percentages, of econometrics-based articles articles in the *JME* and *JMCB*, plus the E-designated articles in the five general-interest journals. The method and data attributes are defined in section 3.2. The 2006–10 figures use data from 2006, 2008 and 2010; and the 2016–18 figures use data from 2016, 2017 and 2018.

Cross-Sectional (Regional) Identification

- A particular useful approach is to explore the cross-sectional/regional exposure to an identified aggregate shock.
- **Example**:
  - Industry-specific international shock (aggregate time-series shock) affects disproportionally places with high-shares of the industry;
  - Monetary policy might affect low-wealth region differently than high-wealth.
- **Issues**:
  - Still need identification (narrative/instrument/diff-diff): is the monetary policy/international shock exogenous?
  - The region-individual must be exposed **before** the shock happen.
  - How to go from cross-sectional estimates to aggregate?

Example: Mian & Sufi (2014, ECTA)

- Impact of "housing net worth" on non-tradable employment at the county-level.
  - weaker household balance sheet ⇒ demand shocks ⇒ decline in real activity;

$$\Delta \log E_i^{NT} = \alpha + \eta \Delta \text{HNW}_i + \varepsilon_i$$

where $\Delta\text{HNW}_i = (\Delta p_{06-09}^{H,i} \times H_{2006}^i)/NW_{2006}^i$.

- **Aggregate Shock**: 2007-2009 housing prices' collapse;
- **Cross-sectional Heterogeneity**: county i housing as a share of net-worth.

Example: Autor, Dorn and Hansen (2013, AER)

- Effect of rising Chinese import competition on US employment.

$$\Delta L_i^M = \alpha + \beta \Delta \mathrm{IPW}_i + \gamma X_i + \varepsilon_i$$

where i denote region and j denote industry:

$$\Delta \mathrm{IPW}_i = \sum_j \frac{L_{ij,1990}}{L_{j,1990}} \frac{\Delta M_j}{L_{i,1990}}.$$

  - **Aggregate Shock**: 1990-2007 change in imports from China by industry (ΔM_j);
  - **Cross-sectional Heterogeneity**: initial differences in regional industry specialization.
  - **Instrumental variables**: Chinese imports by other high-income countries.

Cross-Sectional Identification

- This type of regional identification is known as **Bartik** instrument or **shift-share**. Good references on the econometrics behind:
  - Goldsmith-Pinkham, Sorkin, and Swift (2020, AER), Borusyak, Hull, and Jaravel (2022, ReStud);
- **Main Problem**: in macro sometimes we want to know the aggregate impact of treatment, but the regression only tells you what happens in some regions relative to others;
  - The true aggregate relationship is different than the cross-sectional because of GE effects, spillovers, etc.
  - This is known as **the missing intercept problem**.
  - Example: it could be that trading with China increased aggregate employment, but it did *relatively less* in exposed regions $\Rightarrow$ the regression will say $\uparrow \Delta IPW \rightarrow \downarrow L^M$.

The Aggregation Issue

- **Problem**: (micro-) spillovers?
- **Example**:
  - Bad shock (treatment) in a region induces employment migration to a region that did not receive the shock;
  - Regression identifies differences in employment response from shock between regions $\Rightarrow$ migration increases the difference $\Rightarrow$ estimated β is larger;
  - But aggregate employment did not change as much, it just reallocated from one region to the other.
- **Stable Unit Treatment Value (SUTVA) Assumption**: treatment of one unit does not affect outcomes of non-treated units.
- If this assumption does not hold our regression estimates will be biased.

The Aggregation Issue

- **Problem**: General equilibrium effects (macro spillovers)?
- **Example**:
  - Good shock (treatment) in a region induces more consumption of tradable goods;
  - Price of the tradable good increases;
  - Control regions reduce consumption;
  - **Regression:** good shock increases consumption! but aggregate consumption did not really increase...
- Good theory tells you what the price transmission should look like.

The Aggregation Issue

- **Problem**: Endogenous responses?
- **Example**:
  - Bad output shock (treatment) in a region should not (alone) induce response by the monetary policy;
  - If we are interested in the total impact of the shock in the aggregate, we may want to consider the reaction of the monetary authority...
  - Feature or bug? Maybe we want to know effect of policy or shock separate from effect of policy response.
- Be aware of locally identified tax changes, federal government rebalacing budget, and any interactions between local and federal government.

The Aggregation Issue

- Does that mean we should not use this evidence? NO!
  - The micro evidence can help us to build better models...
  - if the empirical evidence rejects your model go back to the drawing board.
  - Look for the empirical evidence regarding the key elasticities of your model!
- The solutions to the aggregation issues are really problem dependent.
  - Estimate spillovers;
  - Use theory to guide general equilibrium/partial equilibrium effects;
  - Match causal moment in structural model (this is a type of indirect inference);