Many people might give you different definitions of what they think calibration "is" or "should be." These notes will try to understand why this is the case, and also compare to conventional estimation methods.

History The canonical calibration paper was Kydland and Prescott (1982), trying to measure the impact of TFP shocks on business cycles. This old example already gives two perspectives on what a calibration exercise is:

1. They took the standard neoclassical growth model, fed in TFP shocks that match the data, and concluded that TFP shocks explain approximately one third of business cycle fluctuations in the U.S.

2. Alternatively, you could have fed in the TFP shocks so that they exactly match the magnitude of U.S. business cycle fluctuations, but then, without adding more features to the model, would miss out on standard implications from the growth model.

Sadly, things haven't improved so much since then in a philsophical sense: we still debate whether it is best to

1. take empirically valid estimates in a comparable model, feed them into your model and see what you get; or

2. run your model to exactly match what you want to match.

In many cases it's clear which route you want to pursue; in some cases not so much. Calibration began as the former; but recently many papers are closer to the latter. However, the latter must be somehow related to econometrics—but most macroeconomists, for one reason or another, ignore this aspect. This is also a reason we always get hammered by hard-core econometricians.

Today Ever since, there's been a big debate between econometricians and "calibrators" (e.g. the JEP articles in the syllabus). In my opinion, there are two big directions in which calibration has matured[1]:

1. The business cycle literature, whether RBC or New Keynesian, has become more rigorous in a statistical sense (people may disagree with me). It seems the norm today is to do a Bayesian or SMM estimation, unless there is a good reason not to. The empirical counterpart of a business cycle model is more straightforward than a GE heterogeneous agent model in the form of time series.

2. The heterogeneous agent literature relies a lot on "matching moments," rather than estimating. A growing crowd now estimates standard errors assuming an identity weighting matrix, which doesn't really have any econometric foundations. A small crowd does estimate optimal weighting matrices. However, whichever the case, there is a fundamental problem with these models in the sense that a lot of them don't have an empirical counterpart—for example, you can't really estimate an interest rate or wage rate if there is no time variation (in a steady state).

This topics class is mostly about the latter, which I'll focus on in the rest of the notes.

[1] Admittedly, with a heterogeneous agent model with aggregate fluctuations, we could use both time-series and cross-sectional data. However, the method itself is very costly and not recommended unless your focus is on the distributional statistics over the business cycle—therefore, until that literature grows enough to become the norm, I will not consider the problems associated with it.

SMM and Indirect Inference "Two" statistically valid methods that are closely related to calibration, and in which direction I think the literature is evolving, is SMM and Indirect Inference. I say "two" because SMM in a strict sense is a subset of indirect inference. The restriction with implementing indirect inference as opposed to a calibration is that you will have restrictions on which moments you "estimate": you cannot use an interest rate of 4% as an estimated moment if your model implies exactly one draw of a steady state interest rate. You *can* say that you estimate other parameters that *do* have some variation in the data, while also "matching" an interest rate of 4% in a general equilibrium. (As a rule of thumb, no sample or no variation, no identification—or you can say you estimate something with infinite standard errors).

For our problems, SMM is the way to go.

1. Formulation of the Problem

Calibration of heterogeneous agent models in the tradition of Huggett (1993); Aiyagari (1994) usually involve the following nested loops:

1. Choose a parameter vector θ from the parameter space Θ; and fix X, the state space.
   (a) Choose a $x_a \in X_a$.
      i. Assuming aggregate stationarity, find individual decision rules f; typically involves solving a Bellman equation.
      ii. Given (θ, x_a) and f, find stationary distribution Φ.
      iii. Given (θ, x_a) and (f, Φ), check equilibrium condition G.
   (b) Repeat from 1a until 1(a)iii is satisfied.
2. Repeat from 1 until model moments $M(\theta)$ match data moments $M(s)$ from sample s, according to some method.

I now explain each step in detail. Note that, it is *assumed* that individual choices are dynamic; the aggregate state, denoted by X_a, is stationary, and this is common knowledge across all agents:

1. We can split an individual's state space as $X(\theta) = [X_x(\theta)\ X_a]'$, where parts of the state space, X_x, depends on θ (exogenously assumed parametric distributions, stochastic processes, etc.). Hence with some abuse of notation we will suppress the set $X_x(\theta)$ into θ.
2. Solve for individual decision rules $f(x_i; X_a, \theta)$, where $x_i \in [X_x(\theta)\ X_n]'$ are individual states. The rules f can be a vector.
   (a) In most models, x_i (and f) will include savings, and f will include labor supply.
   (b) In any interesting model, there is a relationship $g : f, X_a, \theta \mapsto x_i'$, where the domain of f is implicit for g, but we make X_a explicit to define equilibrium conditions. There may be additional transformations on f according to θ, so we make θ explicit as well.
3. Solve for the stationary distribution of x_i induced by f, call it $\Phi(x_i; f, X_a, \theta)$:

$$\int_{g \in A} g(f, X_a, \theta)\Phi(dx_i) = \int_A x_i' \Phi(dx_i')$$

for any measurable set $A \in \mathcal{B}(X)$. Stokey and Lucas (1989) characterizes when Φ is unique.

4. Given stationarity, a (rational expectations) general equilibrium is found by X_a that solves some conditions $G(\cdot)$ over the decision rules and associated stationary distribution:

$$\hat{X}_a(\theta) \in \{X_a : G(\Phi; f, X_a, \theta) = 0\}, \tag{1}$$

where G depends on X_a implicitly through f, but may also have additional explicit dependencies. Note that even if Φ is uniquely determined by f, we make it explicit in the equilibrium definition, since both GE conditions and moments are expressed over Φ.

5. Moments of interest are now expressed as

$$M(\hat{X}_a(\theta), \theta) = \int_B m(x_i) \Phi \left(dx_i; \hat{X}_a(\theta), \theta \right).$$

for some $m(\cdot)$. Note that $m(\cdot)$ does not depend explicitly on θ. SMM obtains $\hat{\theta}$ that minimizes a weighted mean-squared distance:

$$\hat{\theta} = \arg\min_{\theta} \left[M(\hat{X}_a(\theta), \theta) - M(s) \right]' W(s) \left[M(\hat{X}_a(\theta), \theta) - M(s) \right]$$

for some optimal weighting matrix $W(s)$, which may depend on the sample s.

Note that 5. is a nested fixed point problem. Alternatively we could formulate it as a mathematical problem with equilibrium constraints, following Su and Judd (2012):

5′. Find

$$\begin{aligned} \hat{\theta} &= \arg\min_{\theta} \left[M(\hat{X}_a, \theta) - M(s) \right]' W(s) \left[M(\hat{X}_a, \theta) - M(s) \right] \\ \text{s.t.} \quad & G(\Phi; f, \hat{X}_a, \theta) = 0. \end{aligned}$$

In fact, we can define the SMM estimator in terms of Φ with an additional constraint for stationarity as an equilibrium condition, or even go further and find f by adding yet additional constraints as individual optimality conditions. However, the space of these objects are extremely large and typically cheap to compute using conventional methods—moreover, the above formulation facilitates splitting the parameter space, as in the following subsection.

2. Splitting the State Space

Calibration is an "art" of picking parameters. Now we wish to split the parameter space into 4 regions: those that can be chosen *a priori*, can be chosen independently of the model, are chosen to match macro moments, and finally those that are chosen to match micro moments. Accordingly we will denote them as $\theta = [\theta_0\ \theta_1\ \theta_2\ \theta_3]'$.

1. θ_0 usually consists of preference related parameters of which not much is known, and we choose based on introspection or behavioral evidence, such as risk aversion. Ide-

ally, you do not want to do this unless it is not separately identified from other parameters by construction.[2]

2. Since θ_1 is exogenous to the model, this can be estimated from a data sample without solving the model. For example, in the canonical Aiyagari (1994) model earnings are estimated from the PSID and exogenously fed into the model.

3. (θ_2, θ_3) are more controversial. Obvious candidates for θ_2 are production function parameters in a model with heterogeneous households, or demand parameters for heterogeneous firms, which are chosen to match some time-series average of a moment we can easily solve out for by hand. Then we start to wave our hands and say the rest of θ_2 are chosen in equilibrium while θ_3 are chosen to match cross-sectional moments. However, unless θ_2 is unidentified in a cross-section, we cannot say it is chosen only in equilibrium, since it will affect the micro moments. We will talk more about this below.

   This set will also typically include a policy loop (budget balance, progressive tax rates, etc). See appendix for tax functions used frequently in the literature.

4. Where the literature splits is in how to choose θ_3. As indicated earlier, we can estimate θ_3 from elsewhere, feed it into the model (as we did for θ_1) and see how well it explains other moments of interest. This is calibration in the tradition of Kydland and Prescott (1982). Only do this if you are using a benchmark model where the empirical estimates of a model are well known (Growth Model) and you add a mechanism intended to explain something else (RBC Model). If your goal is to match some moments of interest, the ideal way to go is to estimate it. But how?

3. Research Questions

1. Magnac and Thesmar (2002) show that the discount factor β is generally unidentified in a dynamic discrete choice model without exclusion restrictions. Is this somehow applicable to the Aiyagari (1994) model? If so, we can safely include β as an equilibrium restriction according to (1), and forget about identification. What other moments are naturally unidentified in the cross-section and only found in a general equilibrium?

2. Unlike micro models, heterogeneous agent macro models are not engineered to be naturally mapped into available data sets. Consequently, they refer to many different data sets to pick empirical moments of interest. Is there a consistent (and efficient) way to estimate θ_3 from different data sets? e.g. Petrin (2002)?

3. Are there ways to show identification of a model that rely less on the underlying data structure? e.g., Gentzkow and Shapiro (2013)?

4. Are MPEC packages a reasonable way to numerically implement these models?

[2] For example, Epstein-Zin preferences, which split risk aversion and intertemporal elasticity, are frequently calibrated within the model. This is because these models are engineered to be able to do so.

A. Parametric Tax Functions

Simpler models assume flat tax rates on capital and labor income, that is they assume budget constraints like

$$c + a' = (1 - \tau_l)w\epsilon + (1 - \tau_k)ra + a.$$

Obviously in such cases, we would include (τ_l, τ_k) as exogenous parameters (as a part of θ_0 or θ_1), just assuming the rates are similar in magnitude to aggregate means from the data. More realistically, and in most countries (not all), tax is levied on total income, and typically non-linear (usually it is progressive):

$$c + a' = [1 - t(y)]\, y + a,$$

where $y = w\epsilon + ra$ is the total income flow this period. Guner, Kaygusuz and Ventura (2014) summarize some of the most frequently used, which include:

$$\begin{aligned} \text{Guner et al. (2012)}: &\quad t(y) = a + b\log(y/\bar{y}) \\ \text{Benabou (2002)}: &\quad t(y) = 1 - a(y/\bar{y})^{-1/b} \\ \text{Guvenen et al. (2014)}: &\quad t(y) = a + b(y/\bar{y}) + c(y/\bar{y})^{d} \end{aligned}$$

where $\bar{y}$ is the median or mean income in the economy. This normalization (done in the data pre-estimation; that is, all incomes in a sample are divided by the mean/median income in the sample or some known quantity) is nice, since we can directly use the estimates and just apply the function in the model while also normalizing mean/median income in the model.

The only problem is, we don't know the mean income in the model unless we first simulate it. So we could either

1. Iterate on $\bar{y}$ until the guess of $\bar{y}$ used in the tax function equals the $\bar{y}'$ from simulating the model.

2. Assume a TFP on the aggregate function, which will obviously raise or decrease all individuals' income. Then fix $\bar{y}$ to some constant (typically 1, or just the actually mean value observed in the data, like 30K Euros), and iterate on A until $\bar{y}'$ equals the constant.

In any case, we would be including $\bar{y}$ in θ_2 as an equilibrium parameter. The nice thing of having $\bar{y}$ is that various other policies can also be included without adding extra loops: i) a lump-sum subsidy that you want to be equal to a fraction of mean income, ii) a means-tested subsidy that you only want to give to individuals who earn below $x\%$ of mean income, iii) a capital tax rate (not capital income—they are not the same!) that is only levied on individuals who earn above $x\%$ of mean income, etc. Otherwise we would need to carry multiple policy variables in θ_2, which is computational undesirable.

Lastly, there is the classic parametric tax function (probably the first used in quantitative work):

$$t(y) = a\left[1 - (by^c + 1)^{-1/c}\right],$$

estimated in Gouveia and Strauss (1994, 2000), which is *not* normalized, making it more tedious to use their estimates directly. Besides, their estimates are getting outdated, and the function itself does not jointly capture marginal tax rates for median and high income earners well (actually, neither do Guner et al. (2012)'s or Benabou (2002)'s). Furthermore it has three parameters as opposed to two, while Guvenen et al. (2014) gets top marginal tax rates quite well even assuming $b = 0$ (i.e., with 3 parameters).

The relevant features of these parametric tax functions are summarized in Guner et al. (2014); for more details you can read the original papers.

References

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, August 1994, *109* (3), 659–684.

Benabou, Roland, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, March 2002, *70* (2), 481–517.

Gentzkow, Matthew and Jesse M. Shapiro, "Measuring Sources of Identification in Non-linear Econometric Models," Technical Report, Chicago Booth 2013.

Gouveia, Miguel and Robert P. Strauss, "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, 1994, *47*, 317–339.

_ and _ , "Effective Tax Functions for the U.S. Individual Income Tax: 1966-89," in "Proceedings of the 92nd Annual Conference on Taxation" 2000.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura, "Taxation and Household Labour Supply," *The Review of Economic Studies*, 2012.

_ , _ , and _ , "Income taxation of U.S. households: Facts and parametric estimates," *Review of Economic Dynamics*, 2014, *17* (4), 559 – 581.

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan, "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *The Review of Economic Studies*, 2014, *81* (2), 818–850.

Huggett, Mark, "The Risk Free Rate in Heterogeneous-Agent, Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 1993, *17*, 953–969.

Kydland, Finn E and Edward C Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, November 1982, *50* (6), 1345–70.

Magnac, Thierry and David Thesmar, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, March 2002, *70* (2), 801–816.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, *110* (4), 705–729.

Stokey, Nancy L. and Robert E. Lucas Jr., *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

Su, Che-Lin and Kenneth L. Judd, "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 09 2012, *80* (5), 2213–2230.