Computational Economics

Lecture 10: Using and Estimating DSGE Models

Etienne Briand

Concordia University

Winter 2026

Motivation

In previous lectures we derived and solved DSGE models — to evaluate their usefulness we need to confront them to data.

The fit we get will obviously depend on the structural parameters that we assign which is a non-trivial affair:

- DSGE models are highly parameterized: preferences, technology, and policy each contribute free parameters.
- Many parameters are not directly observable (e.g. habit persistence, Calvo probabilities, adjustment cost curvature).
- The model is an approximation: misspecification is unavoidable, which complicates formal inference.

Overview

We cover four main approaches:

1. Calibration — fix parameters from independent studies such as micro evidence, long-run averages, etc.
2. Moment matching — choose parameters so that model-implied statistics match their data counterparts. Two variants:
    * *Unconditional moments* (variances, auto- cross-correlations, VARs)
    * *Impulse response matching (conditional moments)* à la Christiano, Eichenbaum & Evans (2005)
3. Maximum likelihood — treat the model as a state-space system and maximize the exact likelihood via the Kalman filter.
4. Bayesian estimation — combine the likelihood with prior information over parameters; the dominant approach in modern practice.

Calibration

Calibration

Calibration (Kydland & Prescott 1982) fixes model parameters at values that are:

- Directly identified from *microeconomic evidence* (e.g. labor supply elasticity from micro studies), or
- Implied by *long-run* (steady-state) relationships that are stable and well-measured in the data (e.g. capital-to-output ratio, consumption share).

The logic is one of disciplined parameter assignment: rather than fitting the model to a particular sample, one uses prior economic knowledge to tie down parameters that the time-series data cannot separately identify.

Calibration (cont'd)

Example:

Parameter	Value	Source
Discount factor β	0.99	Annual real rate $\approx$ 4%
Capital share α	0.33	Labor income share
Depreciation δ	0.025	Investment-to-capital ratio
Frisch elasticity φ^{-1}	1 or 2	Micro labor studies
Markup $\theta/(\theta-1)$	1.1–1.2	Industry-level price-cost margins

Limitations:

- Medium-scale models involve many parameters that have no clean micro counterpart — calibration becomes infeasible.
- No formal measure of fit; model evaluation is qualitative.
- Results may be highly sensitive to the parametrization.

Moment Matching

Moment Matching

Moment matching selects the vector of parameters θ to minimize the distance between a vector of data moments $\hat{m}$ and the corresponding model-implied moments $m(\theta)$:

$$\hat{\theta} = \arg\min_{\theta} [\hat{m} - m(\theta)]' W [\hat{m} - m(\theta)]$$

where W is a positive-definite weighting matrix.

The researcher chooses which moments to target (e.g., should the model explain).

Moment Matching (cont'd)

Unconditional moments are statistics computed from data without conditioning on the source of fluctuations:

- Standard deviations: σ_y, σ_π, σ_r, σ_i/σ_y, ...
- Autocorrelations: $\mathrm{Corr}(y_t, y_{t-j})$ for $j = 1, 2, \ldots$
- Cross-correlations: $\mathrm{Corr}(y_t, \pi_{t+j})$, $\mathrm{Corr}(y_t, r_{t+j})$, ...

The model may have closed-form mapping $m(\theta)$ from parameters to these moments via the spectral density of the state-space solution.

Limitation

Unconditional moments conflate *all* shocks. A model that matches second moments can still misrepresent the transmission of any *individual* shock — a feature that monetary policy analysis depends on.

Moment Matching (cont'd)

Christiano, Eichenbaum & Evans (2005) propose to match structural impulse responses from a VAR rather than unconditional moments.

Step 1 — VAR identification: Estimate a VAR and identify a structural shock of interest (e.g. monetary policy shock). Compute empirical IRFs $\hat{\Psi}(\ell)$ with confidence bands.

Step 2 — Model IRFs: For a given θ, simulate the DSGE model's response to the same monetary policy shock: $\Psi(\ell; \theta)$.

Step 3 — Minimize distance:

$$\hat{\theta} = \arg\min_{\theta} \, [\hat{\Psi} - \Psi(\theta)]' W [\hat{\Psi} - \Psi(\theta)]$$

Advantage: The structural shock of interest is explicitly identified; the estimation is transparent and the fit is visually assessable. The fit of the model can be tested on untargeted moments.

Maximum Likelihood Estimation

Maximum Likelihood Estimation

A log-linearized DSGE model can always be written as a linear state-space system:

$$S_t = A(\theta)S_{t-1} + B(\theta)\varepsilon_t \tag{1}$$

$$Y_t = C(\theta)S_t + D(\theta)\eta_t \tag{2}$$

where S_t is the vector of (unobserved) state variables, Y_t is the vector of observables, $\varepsilon_t \sim \mathcal{N}(0, I)$ are the structural shocks, and $\eta_t \sim \mathcal{N}(0, I)$ are measurement errors.

Given this representation, the Kalman filter recursively evaluates the likelihood $\mathcal{L}(\theta \mid Y_{1:T})$ by tracking the distribution of S_t conditional on past observables.

The Kalman Filter: Prediction Step

At each date t, the filter maintains the filtered distribution $S_{t-1} \mid Y_{1:t-1} \sim \mathcal{N}(\hat{S}_{t-1|t-1},\, P_{t-1|t-1})$. It then proceeds in two steps.

Prediction — project forward using the transition equation:

$$\begin{aligned}
\hat{S}_{t|t-1} &= A(\theta)\, \hat{S}_{t-1|t-1} \\
P_{t|t-1} &= A(\theta)\, P_{t-1|t-1}\, A(\theta)' + B(\theta)\, B(\theta)'
\end{aligned}$$

This gives the *prior* distribution $S_t \mid Y_{1:t-1} \sim \mathcal{N}(\hat{S}_{t|t-1},\, P_{t|t-1})$ before observing Y_t.

The prediction error and its conditional variance are:

$$\begin{aligned}
\nu_t &= Y_t - C(\theta)\, \hat{S}_{t|t-1} \\
F_t &= C(\theta)\, P_{t|t-1}\, C(\theta)' + D(\theta)\, D(\theta)'
\end{aligned}$$

The Kalman Filter: Update Step

Update — incorporate the new observation Y_t via Bayes' rule:

$$\begin{aligned} K_t &= P_{t|t-1}\, C(\theta)'\, F_t^{-1} \quad \text{(Kalman gain)} \\ \hat{S}_{t|t} &= \hat{S}_{t|t-1} + K_t\, \nu_t \\ P_{t|t} &= (I - K_t\, C(\theta))\, P_{t|t-1} \end{aligned}$$

Intuitively, the Kalman gain K_t weights how much the state estimate is revised in response to a surprise in the data: large gain when the prior is uncertain ($P_{t|t-1}$ large) or the signal is precise (F_t small).

Initialization

Typically, the filter is initialized at the unconditional mean and covariance of the stationary distribution of S_t, i.e. $\hat{S}_{0|0} = 0$ and $P_{0|0}$ solves the discrete Lyapunov equation $P = APA' + BB'$.

Maximum Likelihood Estimation (cont'd)

Iterating the predict–update cycle over $t = 1, \ldots, T$ yields the sequence of prediction errors $\{\nu_t, F_t\}$. Under Gaussianity, the log-likelihood decomposes as:

$$\log \mathcal{L}(\theta \mid Y_{1:T}) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left[\log|F_t(\theta)| + \nu_t(\theta)' F_t(\theta)^{-1}\nu_t(\theta)\right]$$

MLE then solves:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \log \mathcal{L}(\theta \mid Y_{1:T})$$

Limitations:

- Identification: DSGE likelihoods are often *nearly flat* in certain directions — the filter cannot distinguish some parameter combinations.
- Computational cost: the likelihood surface is non-convex; careful global optimization is required.

Bayesian Estimation

Bayesian Estimation

Bayesian estimation combines the model likelihood with a prior distribution $p(\theta)$ encoding pre-sample information about parameters:

$$p(\theta \mid Y_{1:T}) \propto \mathcal{L}(\theta \mid Y_{1:T}) \cdot p(\theta)$$

The object of interest is the posterior distribution $p(\theta \mid Y_{1:T})$, from which we can extract:

- Posterior mode: $\hat{\theta}_{MAP} = \arg\max_{\theta} \log \mathcal{L}(\theta \mid Y_{1:T}) + \log p(\theta)$
- Posterior mean and credible intervals via MCMC sampling

Why are priors helpful?

Priors act as *soft constraints*: they shrink estimates away from implausible regions of the parameter space where the likelihood is flat, addressing the identification problems that afflict MLE.

Bayesian Estimation (cont'd)

Priors are chosen to:

1. Respect the *domain* of the parameter (e.g. $\zeta_p \in (0,1) \Rightarrow$ Beta prior; $\kappa > 0 \Rightarrow$ Gamma prior).
2. Encode information from *micro studies* or previous estimates.
3. Be *more or less informative*; how much can the data update them.

Bayesian Estimation (cont'd)

The posterior $p(\theta \mid Y_{1:T})$ of DSGE generally does not have a closed form — it is explored via Markov Chain Monte Carlo (MCMC).

Standard algorithm (Metropolis-Hastings):

1. Initialize at the posterior mode $\hat{\theta}_{MAP}$ (found via numerical optimization).
2. At iteration s: draw a *proposal* $\theta^* \sim q(\theta^* \mid \theta^{(s-1)})$ (e.g. a Gaussian random walk with covariance $c \cdot \hat{\Sigma}_{MAP}$).
3. Accept $\theta^{(s)} = \theta^*$ with probability
$$\alpha = \min \left\{ 1, \; \frac{\mathcal{L}(\theta^* \mid Y) \, p(\theta^*)}{\mathcal{L}(\theta^{(s-1)} \mid Y) \, p(\theta^{(s-1)})} \right\}$$
otherwise set $\theta^{(s)} = \theta^{(s-1)}$.
4. Repeat for S draws; discard a burn-in period.

Posterior moments and credible intervals are computed from the retained draws $\{\theta^{(s)}\}$.

Bayesian Estimation (cont'd)

A key by-product of Bayesian estimation is the marginal (data) likelihood:

$$p(Y_{1:T} \mid \mathcal{M}_i) = \int \mathcal{L}(\theta \mid Y_{1:T})\, p(\theta \mid \mathcal{M}_i)\, d\theta$$

Comparing two models $\mathcal{M}_i$ and $\mathcal{M}_j$ via the Bayes factor:

$$\mathcal{B}_{ij} = \frac{p(Y_{1:T} \mid \mathcal{M}_i)}{p(Y_{1:T} \mid \mathcal{M}_j)}$$

$\mathcal{B}_{ij} > 1$ favors $\mathcal{M}_i$. The marginal likelihood automatically penalizes overparameterized models (Bayesian Occam's razor).

In practice the marginal likelihood is approximated by the modified harmonic mean or the Laplace approximation around the posterior mode.

Bayesian Estimation (cont'd)

Advantages:

- Priors regularize estimation and address flat-likelihood identification problems.
- Full posterior quantifies *parameter uncertainty* naturally.
- Marginal likelihood enables coherent model comparison without auxiliary test statistics.
- Now the standard practice at central banks (ECB, Fed, Riksbank, etc.).

Limitations:

- Results can be *prior-sensitive*, especially for weakly identified parameters — robustness checks are essential.
- MCMC is computationally expensive: a medium-scale model requires many ($\sim$100K) Kalman filter evaluations.