# Identification and Estimation

## 1. Quick Summary

| Feature | Identification | Estimation | Calibration |
|---|---|---|---|
| Primary Goal | To determine if model parameters can be uniquely learned from data. | To calculate the "best" numerical values of parameters from a specific dataset. | To assign plausible values to parameters to match observed facts or prior knowledge. |
| Nature | Theoretical | Empirical and Statistical | Pragmatic and often less formal |
| Input | The model's structure and assumptions. | A specific dataset. | Existing studies, micro-data, and stylized facts. |
| Output | A "yes" or "no" answer regarding the possibility of unique parameter values. | Parameter estimates and their standard errors. | A set of parameter values for the model. |
| Stage in modeling | The first conceptual step. | Follows identification | An alternative or complement to estimation. |

---

## 2. What are they?

### 1.1 Identification: The Art of the Possible

It addresses whether it is possible to uniquely determine the true values of a model's parameters from the available data. It's a question of "can we know?" rather than "what is the value?". **Before any data is even collected, an econometrician must first establish if a model is identified.**

- A model is considered **identified** if a unique set of parameter values is consistent with the observed data.
- A model is **unidentified** if multiple, or even infinite, sets of parameter values could have generated the same observed data. In such cases, no amount of data can definitively pinpoint the true parameters.

**Key characteristics of identification:**

- Theoretical: It's a property of the model itself, not the data.
- Prerequisite: It logically precedes estimation. You cannot meaningfully estimate a parameter that is not identified.
- Qualitative: The answer is typically "yes" or "no" (or "partially").

Consider a simple example of supply and demand for a product. If we only observe the equilibrium price and quantity over time, we cannot separately identify the supply and demand curves. Why? Because any shift in either the supply or demand curve will result in a new equilibrium point, and we wouldn't be able to tell which curve (or both) shifted. To identify the model, we would need additional information, such as a variable that shifts one curve but not the other (an instrumental variable).

## 1.2 Estimation: The Science of the Actual

Once a model has been deemed identified, the next step is estimation. This is the process of using actual data to calculate numerical values for the unknown parameters of the model. Estimation provides the "best guess" for the true parameter values based on the information contained in the sample data.

Econometricians use various statistical techniques for estimation, with the choice of method depending on the nature of the model and the data. Some common estimation methods include:

- Ordinary Least Squares (OLS): Used for linear regression models.
- Maximum Likelihood Estimation (MLE): A more general method applicable to a

wide range of models.
- Generalized Method of Moments (GMM): A flexible technique that uses moment conditions of the data.

The result of estimation is a set of estimates (the numerical values of the parameters) and their associated standard errors, which quantify the uncertainty or precision of these estimates.

**Key characteristics of estimation:**

- Empirical: It relies on observed data.
- Quantitative: It produces numerical values for parameters.
- Statistical: It involves statistical inference to assess estimates.

## 1.3 Calibration: The Pragmatic Approach

Calibration is a different approach to assigning values to a model's parameters. Instead of formally estimating all parameters from a given dataset, calibration involves choosing some or all parameter values based on pre-existing knowledge or by ensuring the model's output matches certain key features of the real world.

This pre-existing knowledge can come from:

- Previous studies: Using estimates from other researchers' work.
- Microeconomic data: For example, using household survey data to set a parameter in a macroeconomic model.
- Economic theory: Some parameters might have theoretical justifications for their values.

Calibration is often used in large-scale macroeconomic models, such as Dynamic Stochastic General Equilibrium (DSGE) models, where formally estimating all parameters simultaneously can be computationally intensive or require more data than is available. The goal of calibration is not to test the model in a statistical sense but to create a plausible "laboratory" to conduct policy experiments or understand economic phenomena.

**Key characteristics of calibration:**

- Pragmatic: It often combines information from various sources.

- Non-statistical (in the traditional sense): It does not typically involve formal statistical inference for all parameters.
- Goal-oriented: The choice of parameter values is often driven by the desire to match specific, stylized facts or long-run averages in the data.

# 2. Example Question

Consider the following discrete choice problem. There are two choices $j = 1, 2$. Consumer $i$ has the following utility from each alternative:

$$\begin{aligned} u_{i1} &= \beta x_{i1} + \epsilon_{i1} \\ u_{i2} &= \beta x_{i2} + \epsilon_{i2} \end{aligned} \tag{1}$$

where $\epsilon_{ij} \sim N(\mu_j, \sigma^2)$ for $j = 1, 2$ and they are independent. We observe the covariate $(x_{i1}, x_{i2})$ and the choice $d_{ij} \in \{0, 1\}$ for $i$. Parameters are $(\beta, \mu_1, \mu_2, \sigma^2)$. Which parameters are (and are not) identified in this model?

## 2.1 Solution

- The thing we observe is

$$d_{i1} = \begin{cases} 1 & \text{if } u_{i1} \geq u_{i2} \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

which means that

$$\begin{aligned} Pr(d_{i1} = 1) &= Pr(u_{i1} \geq u_{i2}) \\ &= Pr(\beta(x_{i1} - x_{i2}) > \epsilon_{i2} - \epsilon_{i1}) \end{aligned} \tag{3}$$

- Denote $x_{i1} - x_{i2} = \tilde{x}_i$, and $\epsilon_{i1} - \epsilon_{i2} = v_i \sim N(\mu_1 - \mu_2, 2\sigma^2)$
  Then $Pr(d_{i1} = 1) = Pr(\beta \tilde{x}_i > -v_i) = Pr(v_i > -\beta \tilde{x}_i)$
- Nomarlize it to the standard Normal format:

$$
\begin{aligned}
&Pr(\frac{v_i-(\mu_1-\mu_2)}{\sqrt{2}\sigma} > \frac{-\beta\tilde{x}_i-(\mu_1-\mu_2)}{\sqrt{2}\sigma}) \\
&= 1-\Phi(\frac{-\beta\tilde{x}_i-(\mu_1-\mu_2)}{\sqrt{2}\sigma}) \\
&= 1-(1-\Phi(\frac{\beta\tilde{x}_i+(\mu_1-\mu_2)}{\sqrt{2}\sigma})) \text{ (Symmetry of standard normal)} \\
&= \Phi(\frac{\beta\tilde{x}_i+(\mu_1-\mu_2)}{\sqrt{2}\sigma})
\end{aligned} \tag{4}
$$

- This shows a mapping from model to the data. We can only learn two sets of parameters from the data, or we can only **identify** two sets of parameters: $\theta_1=\frac{\beta}{\sqrt{2}\sigma}, \theta_2=\frac{\mu_1-\mu_2}{\sqrt{2}\sigma}$
- Therefore, $Pr(d_{i1}=1|\tilde{x}_i)=\Phi(\theta_1\tilde{x}_i+\theta_2)$
    - $\theta_1$ is the slope of $\tilde{x}_i$ or marginal effects. **It is identified by the variation across products (characteristics).** To be specific, $x_{i1}-x_{i2}$ identifies it. We get it from choosing different $x_{i1}-x_{i2}$ and see how the choice probability changes.
    - $\theta_2$ is the intercept, or the choice probability when $x_{i1}-x_{i2}=0$. **It is identified by the average choice preference when there is no variation of characteristics.** $Pr(d_{ij}=1|\tilde{x}_i=0)=\Phi(\theta_2)$.
- **We can not separately identify $\beta,\mu_1,\mu_2,\sigma^2$. We can choose to normalize $\sigma^2=1,\mu_2=0$ to identify $\beta,\mu_1$.**