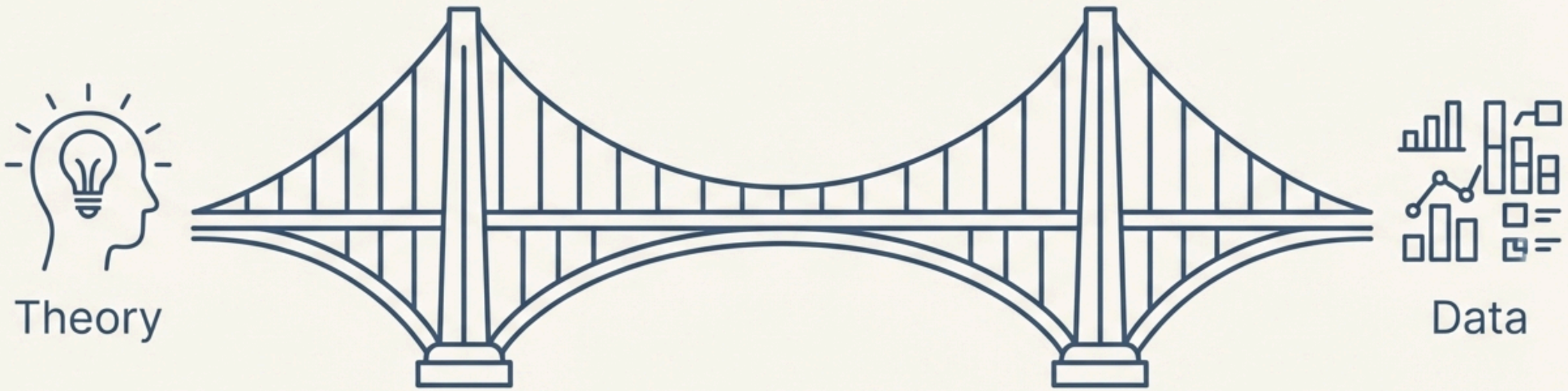
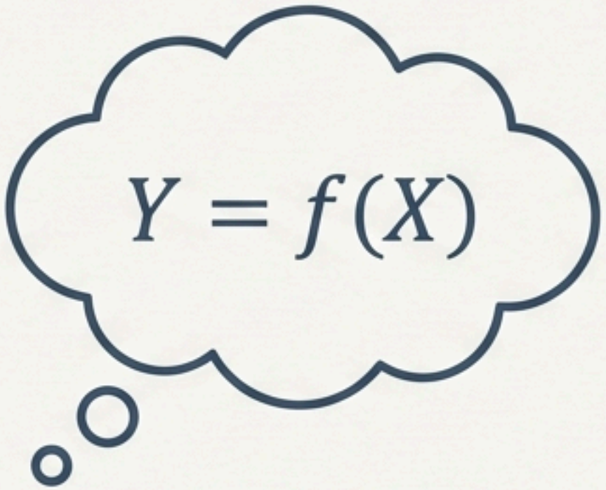
Introduction to Econometrics

Building the Bridge from Economic Theory to Real-World Data



Theories are Ideas. Data is Reality.



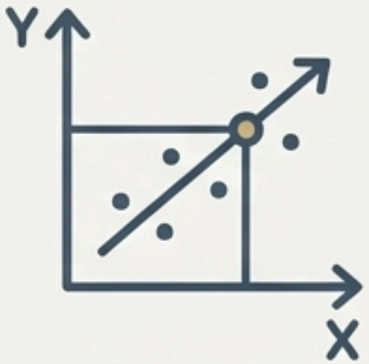
Economic theory provides powerful ideas about how the world works. For example, Keynes's consumption theory suggests a clear link between income and spending.

But the real world is complex and filled with factors that theories don't capture.

How can we use data to test these theories, measure their effects, and make informed predictions?

Econometrics is the Bridge.

Econometrics is the application of statistical methods to economic data. It is the toolkit we use to connect theory with reality.



Estimate Relationships

Quantify the impact of one variable on another.

Test Hypotheses

Determine if real-world data supports or refutes an economic theory.

Forecast Outcomes

Predict future trends based on established relationships.

The 8-Step Econometric Methodology

Econometric analysis follows a systematic and logical process. This roadmap guides us from a theoretical question to a policy-relevant answer.

1. State **Theory** or **Hypothesis**
2. Specify the **Mathematical Model**
3. Specify the **Econometric Model**
4. Obtain the **Data**
5. **Estimate** the Model's **Parameters**
6. **Test Hypotheses**
7. **Forecast** or **Predict**
8. **Use the Model for Control** or **Policy**

Step 1: State the Theory

Our Case Study: The Keynesian Consumption Function

We will walk through the 8-step process using one of the most foundational ideas in macroeconomics.

Keynes's Theory:

As income increases, consumption expenditure also increases, but the increase in consumption will be less than the increase in income.

Core Concept: The Marginal Propensity to Consume (MPC) is positive but less than 1.

Step 2: Specify the Mathematical Model

We translate the verbal theory into a precise, deterministic equation.

$$Y = \beta_1 + \beta_2 X$$

Variable Definitions

Y = Consumption Expenditure (Dependent Variable)

X = Income (Independent Variable)

Parameter Definitions

β_1 = Intercept (Autonomous Consumption)

β_2 = Slope (Marginal Propensity to Consume)

Theoretical Constraint (derived from Step 1): $0 < \beta_2 < 1$

Step 3: Specify the Econometric Model

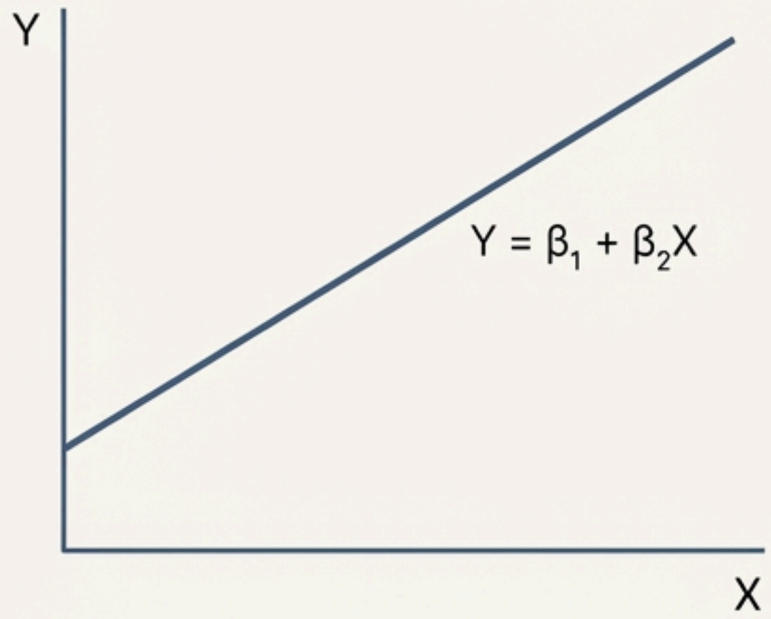
Mathematical models assume a perfect, exact relationship between variables. However, in the real world, relationships are not exact. We account for this randomness, and all other factors that affect consumption besides income, with an error term.

$$Y = \beta_1 + \beta_2 X + u$$

Key Definition: u is the **Disturbance or Error Term.** It is a random (stochastic) variable that captures all factors beyond X that influence Y. It can be positive, negative, or zero.

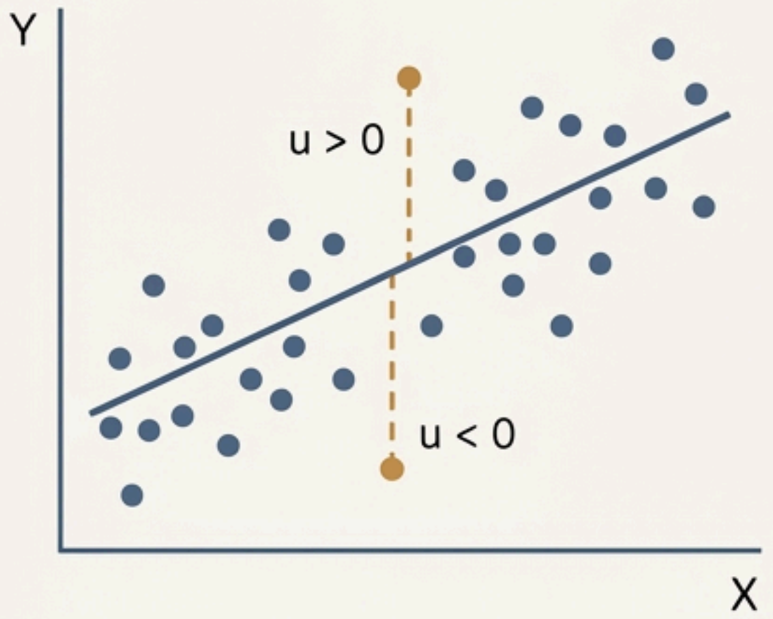
Why the Real World Isn’t a Straight Line

Mathematical Model



The relationship is exact and deterministic.

Econometric Model



The relationship is inexact and stochastic.

Steps 4 & 5: Obtain Data & Estimate Parameters

We use real-world data to calculate the best possible estimates for our unknown parameters, β_1 and β_2. These estimates are denoted as b_1 and b_2.

Sample Data (Unit: Million Baht)		
Year	X: Income	Y: Consumption Expenditure
1980	2,447.1	3,776.3
1981	2,476.9	3,843.1
...	...	...
1991	3,240.8	4,821.0

The "hat" (^) on $\hat{Y}$ signifies that it is an *estimated* value of Y, calculated from our model.

$$\hat{Y} = b_1 + b_2 X$$

Steps 6, 7 & 8: Test, Forecast, and Advise

With the model estimated, we can now use it to generate valuable insights.

Hypothesis Testing (Step 6)

Does our data support Keynes's theory? We test if our estimated parameter b_2 is statistically between 0 and 1.

Forecasting (Step 7)

If we have a projection for income (X) in a future year, we can use our estimated equation to predict the corresponding level of consumption ($\hat{Y}$).

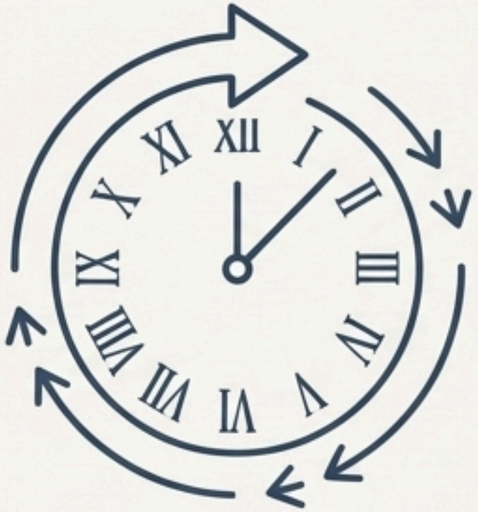
Policy Analysis (Step 8)

Example: If the government wants to achieve a target consumption level of 4,000 million baht for economic stability, what level of national income should its policies aim to generate?

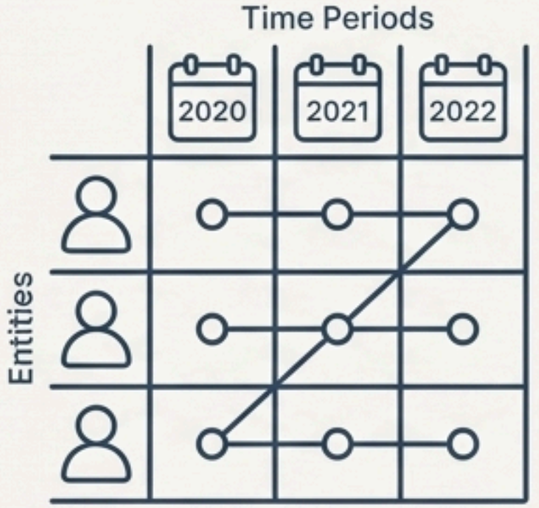
The Foundation of Every Model: Economic Data

The type of data you collect is a fundamental choice in any econometric study. It determines the nature of the questions you can answer and the methods you can use.

Cross-Sectional



Time Series



Panel

Type 1: Cross-Sectional Data

A Snapshot in Time

Definition:

Data collected on one or more variables for multiple entities (like individuals, households, firms, or countries) at a single, specific point in time.

Example:

A survey of income, debt, and education level across 500 different Thai households conducted in the year 2023.

Household Survey (2023)

Household	Gender	Age	Education	Income	Debt
1	Male	62	High School	11,000	0
2	Male	38	High School	15,000	150,000
3	Female	24	Bachelor's	25,000	200,000
...	...	...	...	...	...

Type 2: Time Series Data

A Story Over Time

Definition:

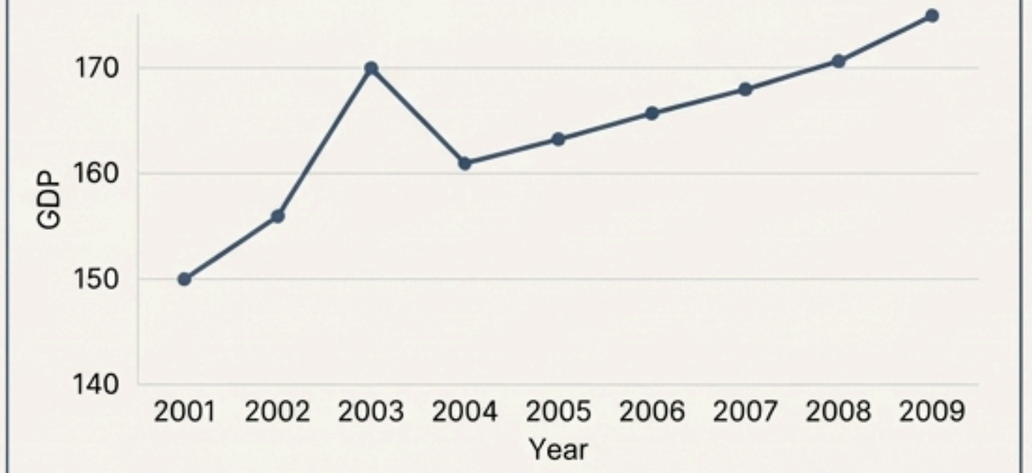
Data collected for a single entity (one country, one company, one market) over multiple, sequential time periods (e.g., daily, monthly, quarterly, or annually).

Example:

Thailand's annual Gross Domestic Product (GDP) from 1980 to 2009.

Thailand Annual Data

Year	GDP	Consumption	CPI
2001	150	82	100.0
2002	156	87	102.0
2003	170	89	107.0
...	...	...	...



Type 3: Panel (or Longitudinal) Data

The Best of Both Worlds

Definition: Data that combines the features of cross-sectional and time series data. The same multiple entities are observed over two or more time periods.

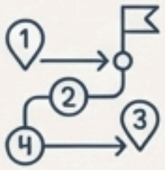
Example: Annual GDP data for Thailand, Malaysia, and Singapore, collected each year from 1980 to 2009.

ASEAN Economic Data

Country	Year	GDP	Consumption	CPI
Malaysia	2002	180	250	118.0
Malaysia	2003	189	257	124.0
Singapore	2002	...	...	...
	2003	...	...	...
Thailand	2002	156	87	102.0
	2003	170	89	107.0

Your Journey as a Modeler Begins

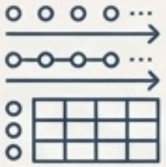
- **Econometrics is the essential bridge** connecting economic theory to real-world data.



- **A systematic 8-step methodology** provides a logical roadmap for analysis.

$\tilde{u}$

- **The error term (u`) is the crucial component** that allows us to model a complex, stochastic world.



- **The choice of data**—Cross-Sectional, Time Series, or Panel—is a foundational decision that shapes your research.

You now have the fundamental framework for building, testing, and applying economic models to understand the world around you.