# 1 Models in Economics

*"Economists do it with models"*

Old Keynesian proverb

A model is a set of equations with economic meaning attached to them. It is important to realize that all models are wrong. Because models are not reality itself, any model we write down will always omit some aspect of reality. Once an element of reality is omitted, it is always possible to find predictions of the model that are at odds with the real world, which means that every model is wrong. Indeed, economic models omit *most* elements of reality, so they are often demonstrably at odds with many empirical observations.

**Even though all models are wrong, some models are useful**. Models have many uses, including differentiating causal from non-causal effects, allowing counterfactual analysis, rationalizing past economic outcomes, providing forecasts, evaluating the social welfare costs and benefits of policy, clarifying assumptions behind economic statements, and giving insight into specific mechanisms otherwise obscured by confounding effects. Many (perhaps most) economists evaluate models based on their usefulness rather than their correctness or accuracy. The choice of what elements from the real world to omit in a model is just as important as the elements that are included. A substantial source of disagreement among economists is, at its core, a disagreement about whether elements of the real world that were omitted by a model are important for the question under consideration and should instead be included.

> **Takeaway**
>
> All models are wrong, some models are useful.

## 1.1 Types of Equations

The models that we study in this course have three types of equations:

- **Identities** hold by definition or construction. For example, disposable income is de-

fined by $Y_D \equiv Y - T$ where $Y$ is income and $T$ represents taxes. When an equation defines a variable, we use the symbol "$\equiv$" rather than a regular equal sign. Another kind of identity is an accounting identity, an equality that must be true regardless of the value of its variables by virtue of how the variables are defined or constructed. The national income identity

$$\text{GDP} = \text{consumption} + \text{investment} + \text{government spending} + \text{exports} - \text{imports},$$

is an accounting identity. Although none of the variables are defined by the accounting identity, the identity always holds (the equality is always true) because of the way in which all the variables are defined. When consumption increases by \$1 and investment, government spending, exports, and imports remain the same, then GDP will always also be measured as having increased by \$1. Another accounting identity is the balance sheet relation that

$$\text{assets} = \text{liabilities} + \text{equity}.$$

- **Behavioral equations** are assumptions that capture some aspect of behavior that we include in a model. For example, a model can postulate that consumers behave so that aggregate consumption is one-half of disposable income. The behavioral equation that captures this behavior is $C = 0.5Y_D$. Behavioral equations hold by assumption and are the equations that distinguish one model from the next.

- **Equilibrium conditions** describe the relationship that must hold between variables in a model when economic forces are balanced. The main equilibrium condition that we will encounter is the condition that supply equals demand. For example, if money demand is assumed to be a function $L(\cdot)$ of the nominal interest rate $i$, and money supply is assumed to be equal to a constant level $M^s$, the equilibrium condition that money supply equals money demand is $L(i) = M^s$.

## 1.2   Types of Variables

All models in economics have two types of variables:

- **Endogenous variables** are variables explained within the model.

- **Exogenous variables** are variables that are taken as given. The behavior and value of exogenous variables are not explained by the model.

Exogenous variables are the inputs that we provide to the model, while endogenous variables are the outputs produced by the model.

The way we use models is by looking at how endogenous variables change when we decide to change the value of one or more exogenous variables. We are not allowed to decide on the value of endogenous variables, or to change their value "by hand" or "by assumption". Only the equations of the model can determine the value of endogenous variables.

Given a set of equations that comprise a model, it is not possible to tell just by looking at the equations which variables are endogenous and which variables are exogenous. A model must include, together with its equations, a list of which variables are endogenous and which variables are exogenous. Whether a variable is endogenous or exogenous is part of the "economic meaning we attach to equations" that was mentioned in the very first sentence of Section 1.

To **solve a model** is to write all endogenous variables only in terms of exogenous variables. To **solve for an endogenous variable** means writing that particular endogenous variable in terms of exogenous variables.

Parameters

We sometimes call certain exogenous variables **structural parameters**, or simply **parameters**. Parameters are exogenous variables that are **policy-independent** or **policy-invariant**, that is, do not change when government policy changes[1]. The exact meaning of policy-invariant can depend on the context or the model. For example, a variable that captures the overall productivity of the workforce is usually considered policy-independent in models of business cycles, but not in models of long-run growth.

---

[1]Economists – myself included! – sometimes consider parameters as a third type of variable (neither exogenous nor endogenous) but I am not sure I can explain why.

Just as for endogenous and exogenous variables, there is no way to know just by looking at the equations of a model which of the exogenous variables are parameters. If a model has parameters, they must be identified separately. Specifying whether a variable is a parameter is not always relevant for the task at hand, in which case we don't need to waste time listing which of the variables are parameters.

*Example* 1.1. We build a simple model of consumption. The model has five variables: consumption denoted by $C$, taxes denoted by $T$, income denoted by $Y$, disposable income denoted by $Y_D$, and the marginal propensity to consume denoted by $c_1$. Taxes, income, and the marginal propensity to consume are exogenous. In addition to being exogenous, the marginal propensity to consume is assumed to be a parameter, whereas income and taxes are not. Consumption and disposable income are endogenous. We assume all variables are strictly positive and, in addition, that $c_1 < 1$. The model has two equations:

$$Y_D \equiv Y - T, \tag{1.1}$$

$$C = c_1 Y_D. \tag{1.2}$$

The first equation is an identity that defines $Y_D$. The second equation is a behavioral equation that describes consumer behavior. Consumers are assumed to consume a fixed fraction $c_1$ of their disposable income. This model has no equilibrium conditions.

Let's now use the model to understand the effect of taxes on consumption. When taxes go up, disposable income goes down by definition. When consumers have a lower disposable income, they reduce consumption.

We cannot use this model to study the effect of consumption on taxes, since consumption is an endogenous variable. It does not make sense to start by saying "consider an increase in consumption" and then trace how the increase in consumption changes taxes. The reason is that taxes, being exogenous, cannot be explained within the model. Any changes in $T$ must be postulated by us; $T$ can only be changed for reasons outside the model. In addition, we cannot simply postulate that consumption increases as consumption is endogenous. If consumption increases, it must be because some change in exogenous variables induced the change.

Now we solve the model, that is, write the two endogenous variables, $C$ and $Y_D$, only in terms of the exogenous variables $Y$, $T$, and $c_1$. Equation (1.1) already expresses $Y_D$ only in terms of $Y$ and $T$, so solving for $Y_D$ is immediate. Plugging equation (1.1) into equation (1.2) gives $C$ only in terms of $Y$, $T$, and $c_1$. Therefore, the model's solution is:

$$Y_D = Y - T,$$
$$C = c_1(Y - T).$$

$\square$

*Example* 1.2. Consider now a different model that has the same equations as the model in the previous example (equations (1.1) and (1.2)), the same five variables ($C$, $T$, $Y$, $Y_D$, $c_1$), but a different set of exogenous and endogenous variables. The exogenous variables of this new model are consumption, income, and the marginal propensity to consume. The endogenous variables are taxes and disposable income.

In this model, it *is* possible to study the effect of consumption on taxes, since consumption is exogenous and taxes are endogenous. Consider then an increase in consumption. Equation (1.2) then implies that the higher consumption leads to higher disposable income. In turn, equation (1.1) shows that, by definition, the higher disposable income is associated with lower taxes.

Let's think about this for a second. This model claims that an increase in consumption leads to an increase in disposable income. Wouldn't that be nice! Of course, this model makes no sense, which is the point: Even though this model has the exact same equations and variables as the model in the previous example, the economic interpretation is completely different because the choice of exogenous variables was different.

One implication is that if in the real world we observe that $C$ and $T$ always go up and down together, two economists using two different models can have a very different interpretation of what is going on in the economy. And they can disagree even if their two models have the exact same equations! $\square$

<u>Shocks</u>

Many times, economists use the word shock to describe changes in exogenous variables. Usually, the term shock is reserved for changes that are unforeseen, unforeseeable, or random, but this is not always the case.

A more formal definition of **shocks** is that they are economically interpretable primitive exogenous forces that satisfy three characteristics:[2]

- They are not explained by past and present endogenous variables.

- They are uncorrelated with other shocks.

- They represent either unanticipated movements in other exogenous variables or news about future movements in exogenous variables.

A technology shock may be the sudden discovery of the steam engine. A tax shock might represent the passing of legislation resulting from a change in political power. The outbreak of COVID-19 can be considered a health shock.

---

**Takeaway**

Models include three types of equations
- Identities hold by construction
- Behavioral equations capture assumptions
- Equilibrium conditions balance economic forces as in supply equals demand

and two types of variables
- Endogenous variables are explained within the model
- Exogenous variables are what the model takes as given
    - Parameters are policy-invariant exogenous variables
    - Shocks are economically interpretable, unanticipated, exogenous variables

---

[2]The definition is adapted from: Ramey, Valerie A. "Macroeconomic shocks and their propagation." Handbook of macroeconomics 2 (2016): 71-162.

## 1.3   Dynamic Models

A model is **static** when it only characterizes endogenous variables at a given point in time. When a model specifies the evolution of endogenous variables over time, we say that it is **dynamic**.

The list of values that a variable takes over time is called a **path**. We use subscripts to index time. When we use **discrete time**, periods are indexed by non-negative integers $(0, 1, 2, ...)$. For example, if $Y_t$ is output in period $t$, then $\{Y_0, Y_1, Y_2\}$ is the path of $Y$ in a three-period model. One period can correspond to a day, a month, a quarter, a year, etc. Suppose GDP in three consecutive years is \$1 trillion, \$1.2 trillion, and \$1.4 trillion. The path of GDP is $\{1, 1.2, 1.4\}$.

We can also use **continuous time** and index time by a positive real number. For example, if time $t$ is a real number in $[0, T]$, then a path for output is the collection of values of $Y_t$ for all $t \in [0, T]$, sometimes also written as $\{Y_t\}_{t \in [0, T]}$ or $\{Y_t\}_{t=0}^{T}$. Just as we had to pick what one period means in discrete time, we have to pick units for continuous time. For example, if one unit of time is a year, our model with $t \in [0, T]$ and $T = 3$ encompasses three years. The value $Y_{0.5}$ is output after six months have passed.

When we specify the values of exogenous variables, we must now specify an entire path for each exogenous variable (one value for each time period) rather than a single value as we did before. Similarly, when we solve the model, we must solve for the entire path of each endogenous variable in terms of the paths of exogenous variables.

---

**Takeaway**

- Static models consider variables at a single point in time
- Dynamic models consider variables that evolve over time
    - The collection of values of a variable over time is called a path
    - Time can be discrete $t = 0, 1, 2, ...$ or continuous $t \in [0, T]$
    - When we solve a dynamic model, we write the paths of endogenous variables in terms of the paths of exogenous variables

---