Chapter 1

Introduction to Dynamic Macroeconometrics

In this chapter, we illustrate several typical properties of dynamic systems by the use of two simple and well-known economic theoretical models: the cobweb model and a textbook macromodel of the Keynesian type. Several concepts which play central roles in the book are illustrated with the aid of the example models. Those concepts include dynamic solution, stability and instability of solutions, equilibrium correction, exogenous variables and dynamic responses to shocks. The chapter ends with an overview of the book.

1.1. Introduction

One hallmark of macroeconomic systems is that the adjustment to a shock (i.e., impulse) spans several time periods. Using terminology invented by the Norwegian economist Ragnar Frisch, whose contributions defined macrodynamics as a branch of economics, we can speak of *impulses* to the macroeconomic system that are *propagated* by the system's internal mechanisms into effects that last for several periods after the occurrence of the shock.[1]

Because dynamic behaviour is an important feature of the real-world macroeconomy as we measure it, a dynamic approach is needed both when the purpose is forecasting, and when the objective is to analyze the effects

[1]Notably, one of Frisch's influential publications was titled "*Propagation Problems and Impulse Problems in Dynamic Economics*" (Frisch, 1933). Frisch's contributions to dynamic theory began even earlier, in Frisch (1929), and in the form of influential lectures in USA (see Bjerkholt and Qin, 2011).

of changes in economic policies. Hence, both fiscal and monetary policies are now guided by dynamic models.

Consider, for example, the use of monetary policy in economic activity regulation, with the central bank sight deposit rate, i.e., the interest rate on banks' deposits in the central bank, as the policy instrument. It is interesting to note that no central bank seems to believe in an immediate and strong effect on the rate of inflation after a change in their interest rate. Instead, because of the many dynamic effects triggered by a change in the interest rate, they prepare themselves to wait a substantial amount of time before the change in the interest rate has a noticeable impact on inflation. The following statement from Norges Bank [The central bank of Norway] represents a typical central bank view:

> The effect of changes in interest rates on inflation occurs with a lag and may vary in intensity. In the time it takes for a change in interest rates to feed through, other factors will also have an impact resulting in changes in inflation and output.[2]

The models used in fiscal and monetary policy analysis need to be quite detailed, and with complex dynamics, in order to meet the requirements of the model users. However, to start studying dynamic econometrics in a systematic way, we can use much simpler models. In this introductory chapter, we look at the *cobweb model*, which illustrates several properties of dynamics that we will meet again later in the book. Section 1.2 presents the cobweb model and its dynamics. In Section 1.3, that model is used to draw the distinction between stable and unstable dynamics, which is important in all system analysis. Two other concepts that are central in the book, *stationary state* and *equilibrium correction*, can also be defined with the use of the cobweb model, see Section 1.4.

The dynamics in Sections 1.2–1.4 is of the deterministic type. However, to make our models applicable to real-world data, we need to include random variables, which we do in Section 1.5, which leads to a simple example of estimation with *time series data* generated by the cobweb model, see Section 1.6.

An attractive property of dynamic models is that they generate the kind of *autocorrelation* that we observe in real-world time series data. Loosely defined, autocorrelated variables are to a certain degree predictable from

[2]Norges Bank's web page on monetary policy. Copied 3rd November, 2016.

their own past. The cobweb model generates the so-called *negative* autocorrelation, while many macroeconomic variables are dominated by *positive* autocorrelation. In Section 1.7, we use a macroeconomic model of the Keynesian type to illustrate positive autocorrelation. At the same time, this provides an example of an open system (of equations), where there are non-modelled, or exogenous, variables.

Section 1.8 gives a brief introduction to the concepts of correlation, regression, exogeneity and causal effect, which are all central in macroeconometrics. The chapter ends with an outline of the content of the book.

1.2. Cobweb Model Dynamics

Our first example of a dynamic model is of partial equilibrium in a single product market. Let P_t denote the price prevailing in the market in time period t, and let Q_t denote equilibrium demand and supply in period t. The model consists of the demand equation (1.1) and the supply equation (1.2):

$$Q_t = aP_t + b, \qquad a < 0, \tag{1.1}$$

$$Q_t = cP_{t-1} + d, \quad c > 0, \tag{1.2}$$

where (1.1) is a static demand relationship: If the price is increased by one (unit), the demand is reduced by an amount a in the same period. Hence, time plays no essential role in this part of the model: the short-run response of demand with respect to a price increase is the same as the long-run response. However, on the supply side, time does play an essential role: if there is a price increase in period t, the supply does not adjust with the amount c before period $t + 1$.

In terms of economic interpretation, equation (1.2) may represent a case of production and delivery lags so that today's supply depends on the price obtained in the previous period. The classic example is from agricultural economics, where the whole supply, for example, of pork or of wheat is replenished from one year to the next.[3]

A different economic interpretation (see, e.g., Evans and Honkapoja, 2001) is that (1.2) captures that the underlying behaviour of suppliers may be influenced by expectations. We then imagine that underlying (1.2) is

[3]For a concise textbook presentation, see Sydsæter *et al.* (2008, Section 11.2).

another relationship, namely,

$$Q_t = cP_t^e + d,$$

where P_t^e denotes the expected price in period t. If we assume that information about period t is either unavailable or is very unreliable in period $t-1$, when expectations are formed, suppliers have to set $P_t^e = P_{t-1}$, which brings us to the dynamic supply function (1.2).

According to Frisch, equation (1.2) clearly qualifies as a *dynamic model* of the supplied quantity since "one and the same equation contains entities that refer to different time periods", namely Q_t and P_{t-1}. What about the model as a whole, i.e., the system given by the static equation (1.1) and the dynamic equation (1.2)? Is it a static or a dynamic *system*? To answer this question, we consider Figure 1.1.

In Figure 1.1, the downward sloping line labelled D represents the (initial) position of the demand equation (1.1). However, care must be taken when interpreting the supply curve. To understand why, consider the quantity supplied in, for example, period $t = 10$ when we assume that $t = 9$ represents a stationary initial state. By stationary (or equilibrium) state, we

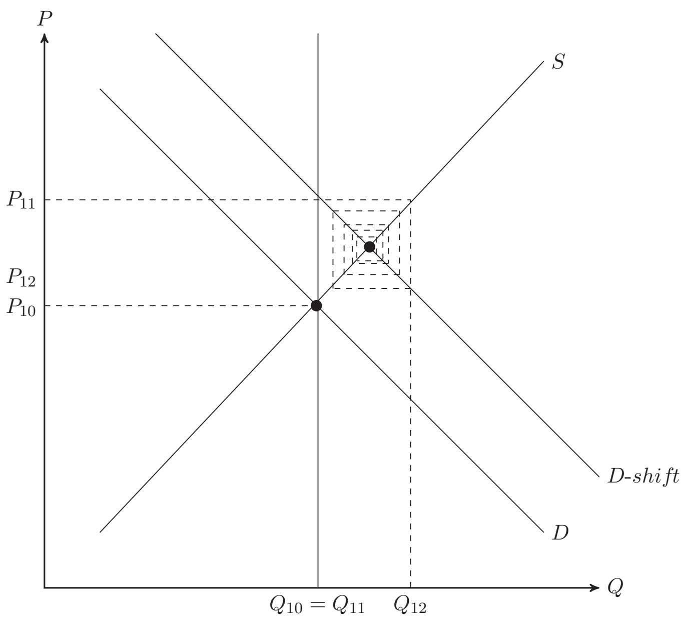

Figure 1.1. Deterministic dynamics: Cobweb model.

mean that both quantity and price have been constant for several periods, hence, for example, $P_9 = P_8 = \cdots = P_{-1}$ and $Q_9 = Q_8 = \cdots = Q_{-1}$.

According to (1.2), supply in period 10 is given by

$$Q_{10} = cP_9 + d,$$

since supply is a function of the predetermined price P_9 (not of the price P_{10}). The supply in period 10, which we can call the short-run supply curve, is therefore completely inelastic with respect to the price in period 10.

The inelastic short-run supply curve is indicated by the vertical line from Q_{10} in the figure. The long-run supply curve is defined for the stationary situation characterized by $P = P_t = P_{t-1}$ and $Q = Q_t = Q_{t-1}$, and is drawn as the upward sloping line labelled S in the figure. Hence, while on the demand side of the market, the derivative of Q with respect to price is the parameter a both in the short-run and in the long-run, there are two different supply derivatives: The short-run derivative is zero, while the long-run derivative is c.

We next turn to the market system's response to a shift in demand which we assume occurs in period 11, represented by the dashed demand schedule labelled *D-shift*. Because of the dynamics of supply, it matters a lot whether the demand shock is permanent or temporary. We first assume that the shock is permanent, and leave the analysis of a temporary shock as an exercise. Even though demand increases in period 11, and because of the inelasticity of supply, the quantity traded in the market in period 11 does not change. Hence, $Q_{11} = Q_{10}$, but the market clearing price increases from P_{10} to P_{11}, i.e., from the old intersection point to the period 1 intersection point between the new demand curve (i.e., *D-shift*) and the short-run supply curve.

The price and quantity pair (P_{11}, Q_{11}) does not, however, represent a new stationary state. Instead, in period 2, the supply will have increased, according to the short-run supply function. With a downward sloping demand curve, P_{12} must therefore be lower than P_{11}. In the example in Figure 1.1, with the slopes being what they are there, P_{12} is nevertheless higher than the initial price, P_{10}. The consequence of the price reduction from period 11 to period 12 is that in period 13, the supply is reduced compared to period 12, and the price rises again. As the figure indicates, the dotted lines that connect the sequence of price and quantity pairs make out a cobweb pattern, and we can imagine that the dynamic adjustment process continues until the new stationary state is reached as indicated by the intersection between the (long run) S-schedule and the *D-shift* line.

1.3. Stable and Unstable Dynamics

With the aid of computer simulation, we can generate data according to the cobweb model, and graph the data in time plots. In Figure 1.2, the two first panels show the time plots of price ("Price dynamics" in the figure) and quantity ("Quantity dynamics") which correspond to the cobweb figure. Initially, the system is assumed to be in a stationary state, and the two plots therefore commence as horizontal lines. In period 11, the demand gets a positive shift, illustrated in the last panel of the figure, and this triggers the dynamic responses in the two first panels. We see that both the price graph and the quantity graph are oscillating in the periods after the demand shift. In time series analysis and econometrics, this kind of erratic behaviour of time series variables is referred to as *negative autocorrelation*.

The solutions shown as time graphs were obtained from the so-called *final form equations* associated with the dynamic model (see Wallis, 1977). A final form equation expresses a current endogenous variable in terms of exogenous variables and lags of itself, but of no other endogenous variable. In Exercise 1.1, you are invited to show that (1.1) and (1.2) imply the two

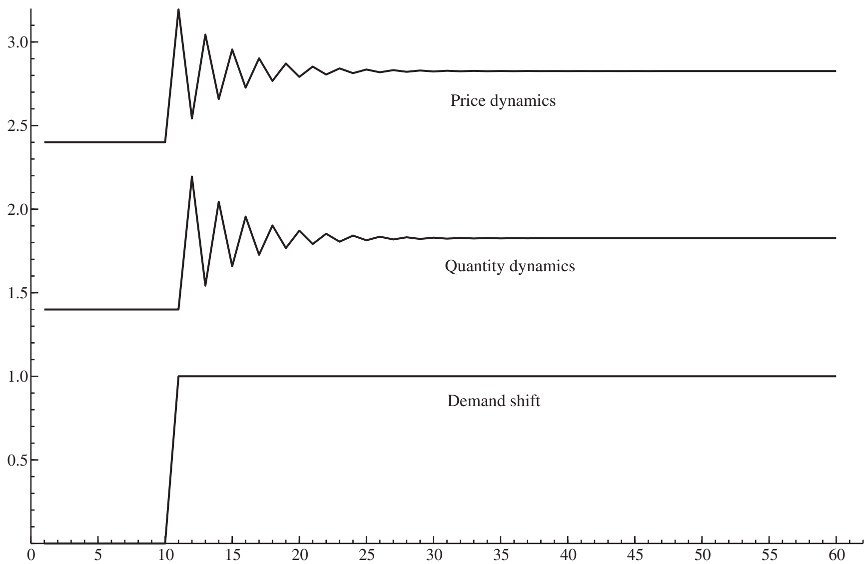

Figure 1.2. Time plots of price and quantity variables that are generated by computer simulation of the cobweb model.

final form equations:

$$P_t = \phi_1 P_{t-1} + \frac{d - b_t}{a}, \tag{1.3}$$

$$Q_t = \phi_1 Q_{t-1} + \frac{da - cb_{t-1}}{a}, \tag{1.4}$$

where the *autoregressive* parameter ϕ_1 is defined as

$$\phi_1 = \frac{c}{a}, \tag{1.5}$$

and b_t, is defined to be b_0 for $t = 1, 2, \ldots, 10$, and $b_1 > b_0$ for $t = 11, 12, \ldots, \infty$.

By looking at the second terms on the right-hand sides of the equality signs, we see that the final equations capture that when b_t shifts, the price is affected first in the same period as the shock. The increase in quantity happens with one period lag due to the delayed response of supply to the market price changes.

In terms of mathematics, equations (1.3) and (1.4) are two first-order linear difference equations (see Sydsæter *et al.*, 2008, Chapter 11). The second terms on the right-hand sign of the two equations are called the *non-homogeneous* parts of the difference equations. Because of the way we have defined b_t, both are functions of time. The first terms in the two equations, $\phi_1 P_{t-1}$ and $\phi_1 Q_{t-1}$, are the *homogeneous* parts of the equations.

Note that the autoregressive coefficient of the homogeneous parts is the same in the two equations. This means that the dynamic properties of P_t and Q_t depend on one single ratio, namely, the ratio of the two slope coefficients, $\frac{c}{a}$. In Chapter 3, where the theory of difference equations is reviewed, it is shown that the condition

$$-1 < \phi_1 < 1 \Leftrightarrow -1 < \frac{c}{a} < 1 \tag{1.6}$$

implies global asymptotic stability, which means that the system consisting of P_t of and Q_t reaches a stationary state from any given initial condition (starting point) as t grows towards infinity.

A stationary state for P_t can be denoted by P^*, and is a parameter that does not depend on time. Stationary states (i.e., equilibria) generally depend on the non-homogeneous part of the final equation, as Figures 1.1 and 1.2 illustrate. In Figure 1.2, P_t and Q_t exhibit dampened oscillations. The oscillations are due to the negative sign of the autoregressive parameter, $\frac{c}{a} < 0$. The dampening is due to the parameter values used in the computer simulation of the model solution, which were chosen to be $c = 1$ and $a = -1.3$. Hence, consistent with Figure 1.1, the supply schedule has a steeper upward slope than the demand curve's negative slope.

There is nothing (in economic theory) hindering that the demand curve has the steeper slope. In that case, with $\phi_1 = \frac{c}{a} < -1$, the system will not be globally asymptotically stable. The dynamic process is unstable, and the oscillations are not dampened; instead, the fluctuations will increase with time as we move away from the period when the shock hit. If $\phi_1 < -1$, the effects of a shock increase with time. Therefore, it is the custom to refer to $\phi_1 < -1$ as the case of *explosive* solution paths. Equation (1.6) implies that $\phi_1 > 1$ also gives rise to explosive solution paths (but without the oscillations). Positive ϕ_1 values are not economically meaningful in the simple cobweb model, but in other models, positive autocorrelation is a typical property. Below, we will look at a macromodel which has $\phi_1 > 0$ as an implication, and where $\phi_1 > 1$ is a possibility.

It remains to be seen what happens to the dynamics of the cobweb model when $\frac{c}{a} = -1$, i.e., the slope of the demand curve is the same as the slope of the long-run supply curve (but of course with opposite signs). In this case, with $\phi_1 - 1$, the dynamic response that follows after a shock is characterized by non-dampened oscillations. Hence, although the dynamic response is not explosive, it is still *unstable*.

The other unstable value that the autoregressive parameter can take is $\phi_1 = 1$, but this value is ruled out by the economic interpretation of the cobweb model. However, for other models, the economically meaningful "candidate" for instability is $\phi_1 = 1$.

1.4. Stationary State and Equilibrium Correction

We can seek a solution where $Q_t = Q^*$ for all t, and $P_t = P^*$ for all t. Such a solution must satisfy $P_{t+1} = P_t = P^*$ and is only possible when $b_t = b_{t-1} = b$. It must also be a solution of (1.1) and (1.2), hence

$$Q^* = aP^* + b,$$

$$Q^* = cP^* + d,$$

which can be solved for the static equilibrium (or stationary) state (Q^*, P^*) as

$$Q^* = \frac{ad - bc}{a - c}, \tag{1.7}$$

$$P^* = \frac{d - b}{a - c}. \tag{1.8}$$

Subject to the stability condition (1.6), we can obtain an equation that shows how $\{Q^*, P^*\}$ are interpretable as attractors to the changes in P_t and Q_t. We use the final equation for P_t, and substitute the expression for ϕ_1:

$$P_t = \frac{c}{a} P_{t-1} + \frac{d - b_t}{a}.$$

Add and subtract P_{t-1} on the right-hand side of the equation:

$$\begin{aligned}
P_t &= P_{t-1} + \left(\frac{c}{a} - 1\right) P_{t-1} + \frac{d - b_t}{a} \\
&= P_{t-1} + \left(\frac{c}{a} - 1\right) \left[P_{t-1} + \frac{1}{\left(\frac{c}{a} - 1\right)} \left(\frac{d - b_t}{a}\right)\right] \\
&= P_{t-1} + \left(\frac{c}{a} - 1\right) \left[P_{t-1} - \frac{a}{a - c} \left(\frac{d - b_t}{a}\right)\right] \\
&= P_{t-1} + \left(\frac{c}{a} - 1\right) \left[P_{t-1} - \left(\frac{d - b_t}{a - c}\right)\right].
\end{aligned} \tag{1.9}$$

Finally, if $b_t = b$ for all t, we can use the definition of P^* to write the final equation as

$$P_t = P_{t-1} + \left(\frac{c}{a} - 1\right) \left[P_{t-1} - P^*\right].$$

By using ΔP_t to represent the change in P_t from period $t-1$ to period t, i.e., $\Delta P_t =: P_t - P_{t-1}$, we can write the final form equation as

$$\Delta P_t = (\phi_1 - 1) \left[P_{t-1} - P^*\right], \tag{1.10}$$

which is an *equilibrium correction model* equation, usually known by the acronym ECM.

Below, we will reserve the term ECM for the case of stable dynamics. In the stable case of $(\phi_1 - 1) > -2$, an initial positive disequilibrium $(P_{t-1} - P^* > 0)$ will lead to $\Delta P_t < 0$. The price fall is larger in magnitude than the deviation from equilibrium, which leads to $P_t - P^* < 0$, and hence $\Delta P_{t+1} > 0$, and so on. The signs of both the disequilibrium terms, and the price changes will change in each period. However, due to the stability assumption, $(\phi_1 - 1) > -2$, the oscillations become dampened over time.

The same steps applied to the final form equation for quantity gives

$$\Delta Q_t = (\phi_1 - 1)\left[Q_{t-1} - \left(\frac{ad - b_{t-1}c}{a - c}\right)\right], \tag{1.11}$$

and, for the case when there is no change in b

$$\Delta Q_t = (\phi_1 - 1)\,[Q_{t-1} - Q^*], \tag{1.12}$$

showing that there is an equilibrium correction equation also for Q_t with the same equilibrium correction coefficient as in the ECM for P_t.

1.5. Random Fluctuations

So far, we have analyzed deterministic dynamics of the type often used in dynamic economics, with the difference that we have used discrete time, whereas continuous time is common in theoretical analysis. However, the hallmark of econometric models is that endogenous economic variables are modelled as random (i.e., stochastic) variables. In the context of the cobweb model, this entails that we introduce two random shocks: ϵ_{dt} (demand) and ϵ_{st} (supply):

$$Q_t = aP_t + b_t + \epsilon_{dt}, \quad a < 0, \tag{1.13}$$

$$Q_t = cP_{t-1} + d + \epsilon_{st}, \quad c > 0. \tag{1.14}$$

The random variables $(\epsilon_{dt}, \epsilon_{st})$ are typically specified with zero means, and with fixed variances. It is the custom to refer to (1.13) and (1.14) as structural equations, and to assume that the two structural disturbances are uncorrelated.

A *reduced form equation* expresses a current endogenous variable in terms of exogenous variables, and of lags of itself and of other endogenous variables. The reduced form of (1.13) and (1.14) is not difficult to find. Equation (1.14) is already a reduced form equation. The other equation is obtained by re-normalizing (1.13) on P_t and then substituting Q_t with the right-hand side of (1.14):

$$P_t = \frac{c}{a}P_{t-1} + \frac{d - b_t}{a} + \frac{\epsilon_{st} - \epsilon_{dt}}{a}, \quad a < 0, \tag{1.15}$$

$$Q_t = cP_{t-1} + d + \epsilon_{st}, \quad c > 0. \tag{1.16}$$

The current period price variable P_t depends on the deterministic demand shock (b_t) and the two current stochastic shocks, ϵ_{st} and ϵ_{dt}. The current quantity, however, only depends on a single within period random shock, namely, the unpredictable supply variation ϵ_{st}. These are the consequences of the adjustment lag in supply, the main economic behavioural assumption of the cobweb model.

The final form equations can be found by following the same steps as for the deterministic version of the model. They become

$$P_t = \phi_1 P_{t-1} + \frac{d - b_t}{a} + \frac{\epsilon_{st} - \epsilon_{dt}}{a}, \tag{1.17}$$

$$Q_t = \phi_1 Q_{t-1} + \frac{da - cb_{t-1}}{a} + \frac{a\epsilon_{st} - c\epsilon_{dt-1}}{a}, \tag{1.18}$$

with notation $\phi_1 = \frac{c}{a}$ for the common autoregressive coefficient in the same way as for the deterministic final form above. In Figure 1.3, we have added the price and quantity dynamics with random-shocks to the graphs with the deterministic solution in Figure 1.2. The stochastic solutions are shown as dashed lines, while the deterministic ones are shown as solid lines. We see that the random shocks add to the volatility of the series. However, the deterministic oscillations dominate in the first periods after the demand shift in period 11. As the system moves away from the period of the demand

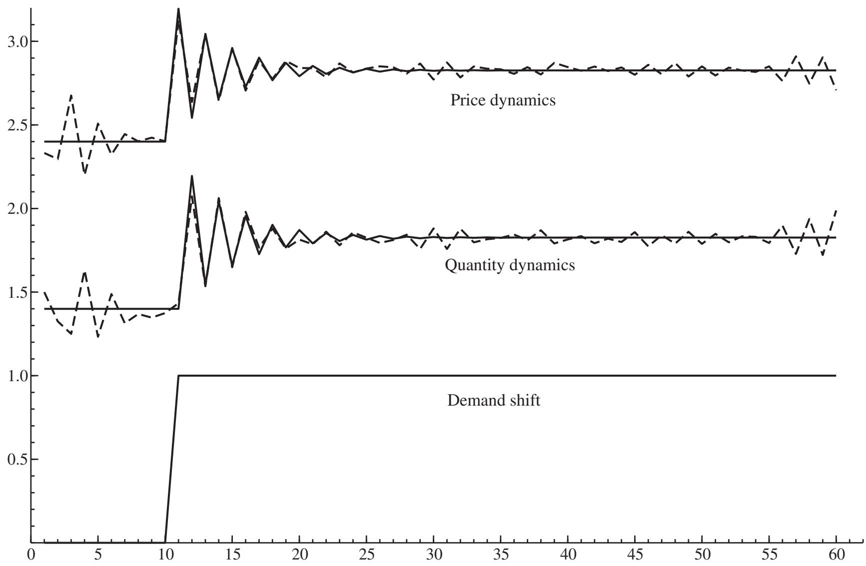

Figure 1.3. Time plots of stochastic and deterministic price and quantity variables generated by computer simulation of the cobweb model.

shock, the effects of the random shocks gradually dominate the fluctuations in the graphs.

1.6. Estimation of Cobweb Model Dynamics

We can use the data plotted in Figure 1.3 to give a first example of the use of ordinary least squares (OLS) to estimate the coefficients of a dynamic model equation. The graph in the first panel of the figure is the time plot of 59 computer-generated observations of P_t, from period 2 to period 60. Assume that we did not know what the value of ϕ_1 used in the data generation was. Can OLS be used to estimate ϕ_1 with a reasonable degree of precision?

To find the answer, we can formulate (1.15) as an regression model equation, for example, as

$$P_t = \beta_0 + \beta_1 P_{t-1} + \beta_2 S11_t + u_t, \tag{1.19}$$

where β_0, β_1 and β_2 are coefficients to be estimated, and u_t is the error term of the regression model. $S11_t$ is a so-called step-dummy (or step-indicator) which captures the shift in the demand function. Hence, $S11_t$ is defined as

$$S11_t = \begin{cases} 1 & \text{if } t \geq 11, \\ 0 & \text{otherwise.} \end{cases} \tag{1.20}$$

OLS estimation of (1.19) by the use of the 59 observations gives

$$P_t = \underset{(0.053)}{-0.755}\, P_{t-1} + \underset{(0.128)}{4.22} + \underset{(0.027)}{0.745}\, S11_t, \tag{1.21}$$

where the numbers in parentheses are standard errors of the estimated coefficients (see Exercise 1.3).

The slope parameter values used in the data generation were $a = -1.3$ and $c = 1$, implying that the true value of the autoregressive parameter is $\phi_1 = -0.769$. Hence, there is an estimation inaccuracy, but it is only minor in this case.

However, it is far from obvious that OLS should work this well. One reason for being sceptical towards the estimation is that the regressor P_{t-1} cannot be exogenous in the sense that you know from introductory econometrics. In the main model taught in introductory courses, often called the "IID cross-section model", the regressor is uncorrelated with *all* error terms, and that assumption drives the result that the OLS estimator is unbiased (see Stock and Watson, 2012, Chapter 14). In the final form equation of P_t, the situation is different. Although P_{t-1} may be uncorrelated with current and future disturbances u_{t+j} $(j = 0, 1, 2, \ldots)$, it cannot possibly be uncorrelated with past disturbances, i.e., u_{t-j}, $j = 0, 1, 2, \ldots$. This follows once we recognize that time flows in one direction from past to present. Therefore, all the shocks that have occurred so far in history must necessarily be "baked into" P_{t-1}, and as a consequence P_{t-1} is correlated with the past regression errors.

The partial dependency between P_{t-1} and the regressions disturbances defines P_{t-1} as a *predetermined variable*. This concept, which characterizes the lagged regressor as a hybrid, partly endogenous, partly exogenous, is central for understanding the possibilities and limitations of estimating dynamic econometric models like (1.21).

Predeterminedness implies that the OLS estimators may be consistent. Consistency of an estimator means that as the number of time periods in the sample increases, the distribution of the estimator gets more and more tightly centred around the true parameter value. Asymptotically (i.e., for an infinitely long sample), OLS estimators are therefore unbiased even though they are biased for any finite sample length. For models like (1.21), the OLS estimators of the coefficients will be consistent if and only if the equation disturbances are not autocorrelated. In the data generation, no autocorrelation was secured by generating the 59 error terms as independent standard normal variables. Hence, a possible explanation of the small estimation error for ϕ_1 that we have observed in the numerical example is that the sample size is large enough to allow the consistency property to "shine through" in the estimation results.

Later in the book, we will review closely the statistical theory that underlies why we have reason to be optimistic about the possibility of estimating the coefficients of dynamic models. Also, why the condition about no autocorrelation in the error term (often called no residual autocorrelation) is crucial.

1.7. Business Cycle Dynamics

The cobweb model's negative autocorrelation, with dominant short-run oscillations for both price and quantity, can be relevant for markets and sectors of the economy where supply is inelastic in the short run. As mentioned above, markets for agricultural products are examples of this, although in modern economies supply can be replenished by trade also for perishable products. A market which has huge macroeconomic influence, and where short-run supply is inelastic, is the housing market. Housing supply is necessarily fixed in the short run (it is a stock), but net housing demand can change a great deal, and rapidly. The implication is that housing prices have a potential for volatility. Another example is the market for foreign exchange under a floating exchange rate regime. Since the net supply of domestic currency is fixed, day-to-day variations in the net demand will determine the exchange rate (Rødseth, 2000, Chapter 1)

However, in markets for manufactured goods and for services, short-run fluctuations in output are mainly demand-driven, and prices follow relatively smooth time paths (due to mark-up price setting). In those parts of the economy, positive autocorrelation in output will be more typical than negative autocorrelation.

As an illustration of positive autocorrelation at the macrolevel, we consider a stylized model of the Keynesian type for a closed economy. The equations of the model are

$$C_t = a + b\,\mathrm{GDP}_t + cC_{t-1} + \epsilon_{Ct}, \tag{1.22}$$

$$\mathrm{GDP}_t = C_t + J_t, \tag{1.23}$$

$$J_t = J^* + \epsilon_{Jt}, \tag{1.24}$$

where the symbols are defined as follows:

- C_t is the private consumption in period t.
- GDP_t and J_t represent gross domestic product, and capital formation (i.e., investment). The three variables are defined in real terms.
- a–c are parameters.
- ϵ_{Ct} and ϵ_{Jt} represent random shocks to consumption and investment with zero means (i.e., mathematical expectation). For simplicity, we assume that they are uncorrelated.

If $c = 0$ is imposed as a parameter restriction in the consumption function (1.22), the model becomes a static macromodel. However, it is reasonable to assume $c > 0$, as it is realistic that private consumption is smoother than income, at least for annual data, which is what we have in mind. The parameter b is the short-run propensity to consume. It is non-negative by assumption, i.e., $b \geq 0$.

Finally, (1.23) is the general budget equation, and (1.24) says, very simply, that J_t, representing investment (capital formation), is fluctuating around the equilibrium value J^*.

The final form equation for private consumption becomes (Exercise 1.4)

$$C_t = \frac{c}{1-b}C_{t-1} + \frac{1}{1-b}\left[a + bJ^* + \epsilon_{Ct} + b\epsilon_{Jt}\right]. \tag{1.25}$$

We see that

$$\phi_1 = \frac{c}{1-b} \tag{1.26}$$

is the autoregressive coefficient in this final equation.

With $c > 0$ and $0 < b < 1$, it is implied that $\phi_1 > 0$, and private consumption is therefore positively autocorrelated.

In the same manner as for negative autocorrelation, positive autocorrelation can be stable, unstable or explosive. As seen from (1.26), stable dynamics requires $c < 1 - b$, which is not implied by the assumptions about consumer behaviour that we stated above. As the cobweb model also illustrated, dynamic stability is a system property, which requires system thinking and system analysis.

If $c \geq 1-b$, dynamics will be unstable ($c = 1-b$), or explosive ($c > 1-b$). With explosive dynamics, we can imagine that consumption is in an initial stationary state, but as soon as there is a small disturbance, consumption will start to move away from that starting value. Unlike the cobweb case, the explosive path will not contain a deterministic oscillating component, but move steadily away from the initial value.

Figure 1.4 shows the time plots of the three variables of the model. The time series were computer-generated, and we have used parameter values $b = 0.25$ and $c = 0.65$, so that the plots illustrate stable dynamics. We see that the volatility of the investments series is clearly reflected in the time plot of GDP. In fact, since J_t is autonomous and fluctuates randomly above and below $J^* = 40$, the only "source of" positive autocorrelation in the two simultaneously determined variables GDP_t and C_t is the smoothing of consumption, which is clearly visible in the time plot of C_t.

Returning to the theoretical model, the stationary state is obtained by solving (1.22)–(1.24) for the case where all shocks are set to zero and $C_t = C_{t-1} = C^*$. Hence, we find expressions for C^*, GDP^* and J^* from the static equation system:

$$C^* = a + b\,\text{GDP}^* + cC^*, \tag{1.27}$$

$$\text{GDP}^* = C^* + J^*, \tag{1.28}$$

$$J = J^*. \tag{1.29}$$

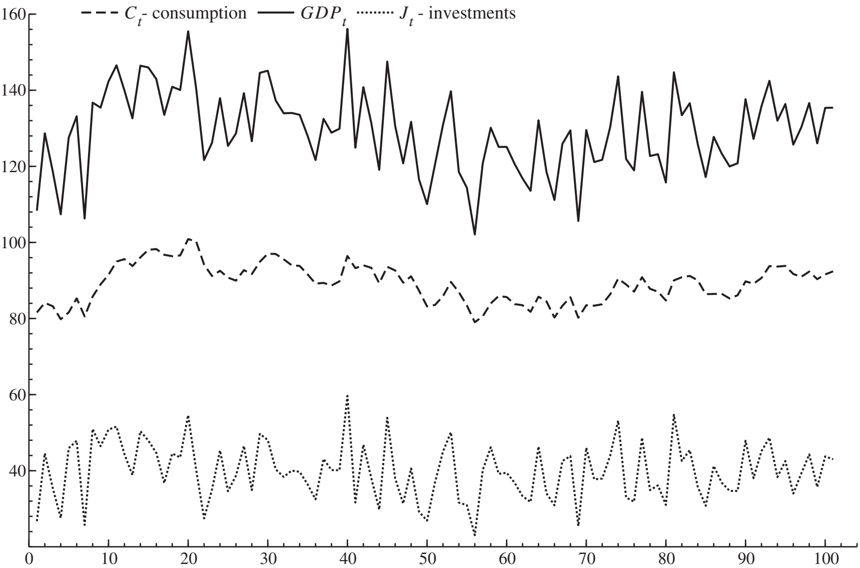

Figure 1.4. Time plots of computer-generated data for the variables of the Keynesian model (1.22)–(1.24).

The solution for C^* can be expressed as

$$C^* = \frac{a + bJ^*}{(1 - b - c)}. \tag{1.30}$$

This solution is only meaningful when the denominator is strictly positive, which is implied by positive and stable autocorrelation:

$$0 < \phi_1 < 1 \Longrightarrow 0 < 1 - b - c < 1. \tag{1.31}$$

In the case of stable dynamics, the final form equation for C_t can be reexpressed as

$$\Delta C_t = (\phi_1 - 1)\,[C_{t-1} - C^*] + \frac{\epsilon_{Ct} + b\epsilon_{Jt}}{1 - b}, \tag{1.32}$$

which has an equilibrium correction interpretation, but without the "overshooting" that characterized the cobweb model: Since $(\phi_1 - 1)$ is positive and less than one in absolute value, the correction taking place in period t is less than the disequilibrium in period $t - 1$.

Box 1.1. RBC and autocorrelated output-gap

Other macrodynamic models can have completely different theoretical interpretations than the Keynesian. Appendix A contains a Real Business Cycle (RBC) model, and a simulated series for the (log of) GDP minus trend, the so-called output-gap.

As noted, in the generation of the time series plotted in Figure 1.4, we have chosen parameters $b = 0.25$ (short-run marginal propensity to consume), and $c = 0.65$ (habit formation). The calibrated autoregressive coefficient ϕ_1 is therefore 0.714, meaning that on average one-third of the disequilibrium from last period becomes corrected in the current time period.

1.8. Causality, Correlation and Invariance

Some introductory books in econometrics liken econometric estimation with laboratory studies where the focus is on how a treatment, measured by variation in a variable X, causes an effect or response, in a variable Y. The research interest is then on the causal mechanism, from X to Y: However, unless the variation in X is carefully controlled in a laboratory experiment, the estimation of the causal mechanism needs to be based on much more than "mere" data and statistical methods. Specifically, as the reader may be well aware of, correlation alone cannot distinguish between cause and effect. With reference to Figure 1.5, if we imagine a dataset with observations of Y and X, and that we calculate the squared empirical correlation coefficient r^2_{YX}, that number can be close to 1, but that does not imply that X caused Y. Instead, Y may have caused X as illustrated in Figure 1.6. The correlation coefficient r_{YX} is the same in the two cases. However, the slope coefficients of the two regressions corresponding to Figures 1.5 and 1.6 are not identical.

By solving Exercise 1.6, you are reminded that for the two-variable systems in Figures 1.5 and 1.6, the relationship between the two slope

Figure 1.5. Two variable system with (one-way) causation.

Figure 1.6. Two-variable system with reverse causation. r_{YX} is invariant to the direction of causation.

coefficients that we can estimate is

$$b = b' \frac{s_X^2}{s_Y^2}, \tag{1.33}$$

where b corresponds to Figure 1.5 and b' corresponds to Figure 1.6. s_Y and s_X are the empirical (sample) standard error deviations (i.e., $s_Y = \sqrt{s_Y^2} = \sqrt{1/T \sum_{t=1}^{T} (Y_i - \overline{Y})^2}$, for Y). We use the symbol T for the sample size (number of periods).

By itself, the fact that the estimated numbers for b and b' are different is of little help when it comes to settling the causality issue. However, it seems reasonable to require of our tentative causal relation that it remains more or less unchanged when other parts of the system undergo changes. Without that kind of *invariance*, the relationship will be of little value when we attempt to use it to predict the effects of changes, see Box 1.2.

Box 1.2. Invariance and persistence of economic relationships
The importance of judging the degree of persistence over time, and the invariance with respect to shocks, of relationships between economic variables, was one of the many fundamental issues that were raised and tackled, in principle, in the treatise "The Probability Approach in Econometrics", by Haavelmo (1944); see Morgan (1990), Bjerkholt (2007), Hendry (2018) among others. Chapter 2 of the treaty was titled "The degree of permanence of economic laws", which also introduced the term *autonomy* in this context. Much later, invariance and autonomy became integrated in the conceptual and theoretical developments related to the role of exogeneity in econometric models, see Chapter 8.

With the invariance requisite brought into the picture, equation (1.33) points to one interesting possibility: namely that at most one of the two regression coefficients is stable (invariant to shocks), when the ratio $\frac{s_X^2}{s_Y^2}$ changes. If b is invariant in this way, b' cannot be invariant, and vice versa. Hence, invariance can be investigated empirically over a sample that contains shocks. This is at least something: If invariance in "one direction" can be established over a sample where there has been interventions (i.e., policy changes), there is some reason to hope that the invariance may also hold for the next change that we want to estimate the effect of. Realistically speaking, relationships between economic variables cannot be invariant to all types of shocks. Invariance is a partial and relative property. Econometric relationships and models are products of civilization, and as such, they will be disrupted sooner or later. Nevertheless, invariance (partial as it may be) is a valuable property of an econometric relationship.

Correlation, regression and causation have been on the minds of statisticians, econometricians and philosophers since the start of the previous century. The initial position may have been that causality *was* correlation. But the statisticians Karl Pearson and George Udny Yule showed that there were important exceptions. The most important ones are called *nonsense regression* and *spurious correlation*. Yule's (1926) nonsense regression has become better known among time series econometricians as *spurious regression* following the influential work of Granger and Newbold (1974). It denotes the case where standard tests for the existence of a relationships are used in data situations where they do not apply. As a consequence, the tests may strongly indicate a relationship which is not in the data. Spurious regression is a main theme in Chapters 9 and 10.

While the spurious regression phenomenon is due to a serious underestimation of the probability of Type-I error when testing the null hypothesis of no-relationship, spurious correlation is about interpreting a relationship that does in fact exist. Spurious correlation can be explained by comparing Figure 1.5 with Figure 1.7, where we have introduced a third variable, Z. If the reality is that both X and Y are influenced by Z, the correlation

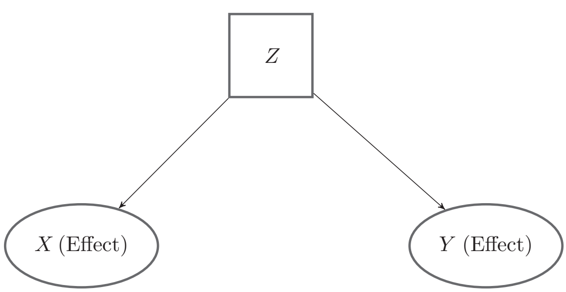

Figure 1.7. Spurious correlation: Two variables with one common cause: omission of Z may lead to wrong conclusion about causation between Y and X.

between X and Y in Figure 1.5 can be called spurious: that correlation reflects that the two variables have a common cause, not that there exists an independent relationship between them. When the situation is as shown in Figure 1.7, it leads nowhere to try to affect the Y variable by taking control over X. Instead, the key is to model the relationship between Z and Y, to become able to use it to make predictions about how Y responds to a change in Z.

Box 1.3. Find out more about causality and exogenity
Aldrich (1995) is a very readable journal article about Pearson's and Yules' contributions to correlation analysis. Causality in macroeconomics is treated in depth in the book by Hoover (2001). A modern philosopher who has written about causality and exogeneity in economics is Cartwright (1999, 2007). Judea Pearl is a central figure in the causality debate, see his "Causality blog": http://causality.cs.ucla.edu/blog. Exogeneity is the topic of Chapter 8 below.

Later in the book, we will present statistical methods that allow us to analyze the degree of invariance of empirical relationships, and in some cases the methodology makes us able to reach a conclusion about the direction of causality within the realms of the models that we are investigating. However, in macroeconomics, it is just as important to be able to analyze systems with two-way dependence (i.e., two-way causality), as depicted in Figure 1.8. Both the cobweb model and the macromodel above are examples of joint dependency between endogenous variables, and of dynamic forms of interdependency. For example, the dynamic responses to a demand shift

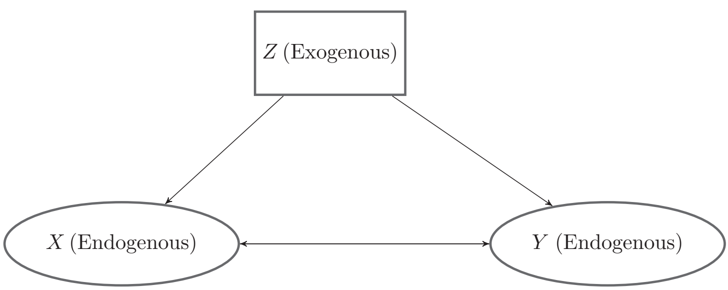

Figure 1.8. Mutual dependence.

in the cobweb model, which we discussed at the start of the chapter, can be interpreted as an example of change in the Z variable in Figure 1.8.

The Keynesian macroeconomic model above is another example of joint dependence since income depends on consumption and consumption depends on income in the same period. Hence, the joint dependence is represented by a simultaneous equations model (SEM). The exogenous variable J_t in the model is represented by Z in Figure 1.8.

In the same way as the term "cause" can take different meanings, and as the conditions for causal analysis can hold under some changes to the system (e.g., policies) and fail under others, there are several concepts and connotations of the meaning of "exogenous variable". When we study macroeconomic theory, "being exogenous" means that the variable in question is determined outside the model. The model is open in the sense that the number of equations is fewer than the number of variables.

Also, in econometrics, the concept of exogenous variable sometimes refers to being determined outside the model, but for some purposes we need to be more precise by using a relevant concept from a wider family of econometric exogeneity concepts, as explained in Chapter 8. Underlying these concepts is, however, the idea that a variable may be regarded as exogenous if we do not distort the answer to the research question we have posed by conditioning on that variable.

1.9. Overview of the Book

Econometrics as a field of knowledge is a combined discipline, as illustrated in Figure 1.9. It combines economic theory with mathematical statistics. Empirical econometric models in particular not only require data for specification and testing, but both producers and users of econometric models need to have knowledge about data sources, and of the main strengths and limitations of the measurement system.

With the exception of Chapter 2, where we review econometric background about models and estimation method for *cross-section data*, the data type that we have in mind in the book is *time series data*, where the variables are observed for several time periods. However, there are several concepts and methods that we cover, which may be of interest for researches who estimate models with the use of *panel data*, which combine cross-section variation and time series variation.

The statistical theory of stationary time series is well developed and is the foundation of applied time series analysis. It also gives the basis of

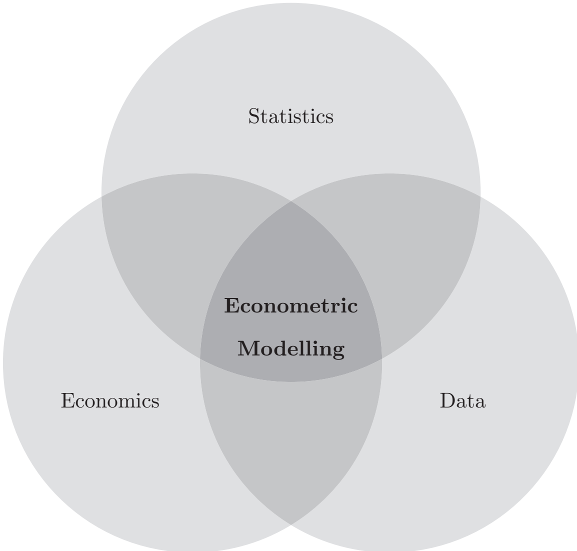

Figure 1.9. Illustration of Econometric Modelling as the intersection of the fields of statistical theory, economic theory and the information in observed data.

econometric models for stationary time series data. This basic theory is presented in Chapters 3 and 4. In Figure 1.9, it can be represented by the intersection between the area labelled Statistics, and the area labelled Economics. Applications of these models to real-world data, we can imagine as belonging to the intersection area labelled Econometric Modelling, where statistics, economics and data combine to become empirical econometric modelling.

Chapter 2 contains a review of the concepts and theory from statistics and econometrics that we assume that the reader is familiar with. That chapter covers, in a concise way, regression models and the estimation of structural equations estimated by instrumental variables. It also introduces matrix notation for econometric models. Chapter 3 is also a background chapter with a presentation of the necessary theory of linear difference equations.

Chapter 4 defines a stationary times series variable as a stochastic process which is mathematically expressed by a linear difference equation which is globally asymptotically stable. Although much of the basic concepts and theory can be learned from studying the simplest case, a first-order autoregressive model, AR(1), applied modelling requires higher order dynamics.

Chapters 5–7 cover the statistical and econometric analysis of stationary economic systems. These chapters contain the core econometric tools and results for building empirical econometric models of stationary systems of time series variables, both single equation models and multiple equation models. The vector autoregressive (VAR) system in Chapter 5 is both an important model in itself, and a basis for the econometric models in Chapter 6 (single equation models) and Chapter 7 (multiple equation models).

In Chapter 8, several important concepts of exogeneity are defined. Weak exogeneity is perhaps the fundamental concept, since without it, consistent estimation of parameters of interest is impossible. Most often weak exogeneity, characterizes a valid conditioning variable in the regression model, but can also refer to a instrumental variable, and to a non-modelled variable in a multiple equation econometric model. Other concepts are relevant when the research objective is forecasting and policy analysis. Invariance, which was mentioned above, is a concept which is closely connected to exogeneity.

In Chapter 9, the theory is extended to non-stationary time series. There are two broad forms of non-stationarity, deterministic trends and unit-root non-stationarity. Deterministic non-stationarity includes, in addition to the (linear) time trend t, t^2 and higher order polynomials in t, and impulse indicator variables and step-indicator variables. Deterministic non-stationarity can be incorporated in the stationary framework of the earlier chapters, and, in practice, statistical inference is unaffected by this kind of non stationarity.

Non-stationarity due to unit-roots is another matter. Models with unit-roots are important since they are able to reproduce typical properties of real-world time series. However, the statistical inference theory is affected, notably when we test for relationships between independent unit-root non-stationary variables, i.e., the spurious regression problem which was briefly noted above.

Chapter 10 is about cointegration which can be regarded as the "flip of the coin" of spurious regression. Unit-roots make inference about relationships between variables hazardous when standard critical values are used, but the problem is greatly reduced by using modern, non-standard, distribution theory. Series that are cointegrated have several representations that are known from the stationary case, notably equilibrium correction, VAR and moving average. And vice versa, equilibrium correction implies cointegration. The VAR representation of cointegrated variables is known as a cointegrated VAR. Dynamic models of the cointegrated VAR can be developed, and represent a major increase in the scope of econometric modelling of stationary systems.

Empirical model specification involves many choices and decisions, for example, about sample length, variable transformations, order of dynamics, significance levels, small or large final model to report, etc. If we can use computer algorithms to automatize some of these many decisions, we can speak about semi-automatic modelling in econometrics (Hendry, 2018).

A long-lasting debate in econometrics is whether it is best to go specific-to-general, or whether general-to-specific is to be preferred. In Specific-to-general: start from small model, and enlarge if it fails. In General-to-specific (Gets): start from a large model, and reduce it to a smaller one with the aid of statistical tests. Both types of searches use economic theory and both make use of several statistical tests, and practitioners use both.

In Chapter 11, we discuss automatization of modelling decisions with the use of the algorithm *Autometrics*, which is part of the program package PcGive that we refer to in the book, (see Doornik, 2009; Doornik and Hendry, 2013a). The main purpose of automatic variable selection is to find the data generating process (DGP). Another usage, perhaps less obvious at first thought, is to robustify a theoretically specified model against the biases that will otherwise affect the parameter estimates, i.e., unless the *a priori* specification should happen to be identical to the DGP. In Chapter 11, we provide examples of both usages.

Forecasting is an important purpose of statistical econometric model building. Chapter 12 discusses the role of empirical macroeconometric models in forecasting. There are several good reasons for doing model-based forecasting, especially when the number of variables forecasted is relatively large. For example, model-based forcasts obey the accounting identities that are specified in the macroeconometric model. Macroeconomic forecasts of cointegrated variables also are themselves cointegrated. None of this "internal order" on a set of forecasted variables is possible by relying on assembling statistically sophisticated univariate forecasts.

Nevertheless, model-based forecasts frequently experience failures in the sense that the forecast errors are larger than there was reason to believe on the basis of the model's performance within sample. The chapter therefore aims to elucidate the weak spots of model-based forecasting as much as its strengths. As we shall see, the same features that make a model attractive, such as economically interpretable long-run relationships and equilibrating adjustment, make the model-based forecasts vulnerable to *structural breaks* in the economy. Hence, it is not only model specification that requires the active decisions of the model builder, but the use of the model in forecasting also depends on the intervention of the model user to aid the model's adaption to a new reality after a structural break in the economy has occurred.

Figure 1.10 shows the chapters that we have briefly reviewed, and some of the relationships between them. The solid lines indicate the "route map" that simply follows the numbering of the chapters. The dashed lines indicate alternative routes through the book.

For readers who have background in an introductory course in econometrics, which covered both regression models (estimated by OLS) and structural model equations (estimated by Method of Moments/Instrumental Variables), the material in Chapter 2 may be well known. The use of matrices to represent those model equation types may, however, not be covered by an introductory course. In any case, it is definitively possible for readers with a good grounding in econometrics to jump to Chapter 3.

Often, a first econometrics course gives an introduction to dynamics, for example, under the heading dynamic regression (models). For readers with that kind of background, it may be possible to skip Chapter 3, and move directly to Chapter 4 (as indicated by the dashed line) where the concept of stationary time series is defined, and the properties of stationary time series are presented. However, Chapter 4 (and indeed the rest of the book) makes many references to mathematical concepts and results in Chapter 3, so the reader should at least be prepared to use Chapter 3 as a reference.

As noted above, we view both single equation econometric models (Chapter 6) and multiple equation models (Chapter 7) as models of the VAR (Chapter 5). It is possible to pass from Chapter 5 directly to Chapter 7, but note that one important class of multiple equation models consists of conditional model equations and marginal model equations, hence there is a close connection between Chapter 6 and Chapter 7.

On the other hand, it is possible to go directly from Chapter 6 to Chapter 8 (exogeneity) since the main discussion in Chapter 8 relates to parameters of interest in single equations models. However, as explained in the chapter, exogeneity is also relevant for multiple equation models.

The discussion of Gets modelling and automatic variable selection (Chapter 11) refers mainly to regression models, so it can be read after Chapter 6 as indicated. However, the chapter also refers to non-stationarity

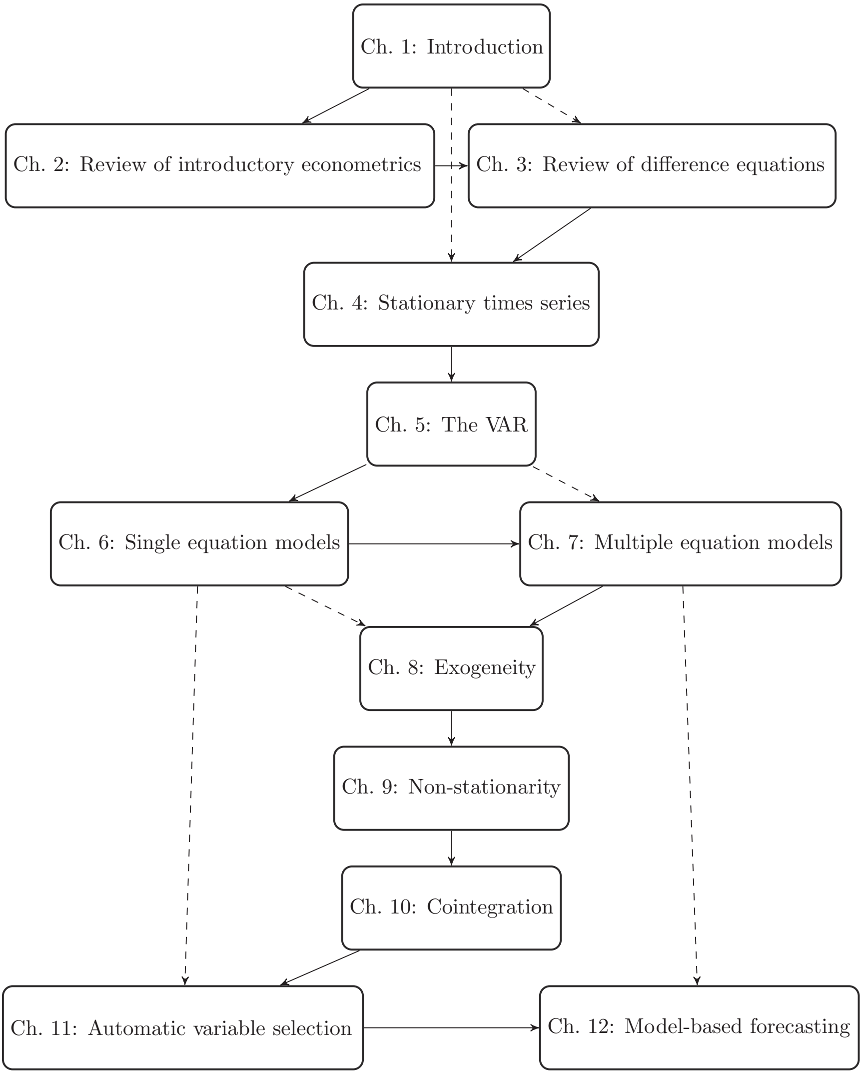

Figure 1.10. The main chapters of the book, with some of alternative "route maps" indicated by dashed lines.

and cointegration, so an even better choice is to simply follow the ordering of the chapters. The last chapter of the book discusses forecasting from a system perspective. Since the exposition makes use of the simplification that the model parameters are known within the sample (estimation theory is tuned down), the chapter can be read, for example, after having been through Chapter 7), as indicated.

1.10. Questions and Answers and Statistical Software and Datasets

There are exercises at the end of each chapter, and Appendix C contains answer suggestions to all the exercises. The questions are a mix of theoretically exercises, and computer exercises.

The book mainly makes use of the *OxMetrics* software programs *PcGive* and *PcNaive*, which have been developed by Jurgen A. Doornik and David F. Hendry at the University of Oxford over a number of years. Information about the latest version of the *OxMetrics* program family is available at: www.timberlake.co.uk/software/oxmetrics.html.

In addition, *Eviews* is used in the solution sketches to a selection of the computer exercises to demonstrate the replication of results from different programs, and to show that most of the exercises do not require a specialized software. Eviews is at www.eviews.com/home.html.

The datasets used in the examples, and in the exercises (at the end of each chapter) can be downloaded from the book's website: http://normetrics.no/dynamic-econometrics-for-empirical-macroeconomic-modelling/

1.11. Exercises

Exercise 1.1. Show that (1.3) and (1.4) are the final form equations for the system

$$\begin{aligned} Q_t &= aP_t + b_t, \qquad a < 0, \\ Q_t &= cP_{t-1} + d, \quad c > 0. \end{aligned}$$

Exercise 1.2. Explain how Figure 1.1 is affected if the assumption $-1 < \frac{c}{a} < 0$ is changed to $\frac{c}{a} = -1$.

Exercise 1.3.

(1) Estimate (1.19), using the dataset *dgp_demandandstoch* in one of the formats provided on the book's webpage. Use the variable labelled P_STOCH as the price P_t, and the series named STEP11 in the database as $S11_t$ in the text.

(2) Estimate the model using the first 19 time periods. What are the main differences compared to the full sample results?

Exercise 1.4. Derive the final form equation (1.25) from models (1.22)–(1.24).

Exercise 1.5. Use models (1.27)–(1.29) to derive the expression in (1.30), and a solution for GDP*.

Exercise 1.6. Show that for the 2-variable systems in Figure 1.5 and Figure 1.6, the relationship between the two slope coefficients that we can estimate is

$$b = b' \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2},$$

where b corresponds to the relationship shown in Figure 1.5, and b' corresponds to Figure 1.6. $\hat{\sigma}_Y$ and $\hat{\sigma}_X$ are empirical (sample) standard error errors: $\hat{\sigma}_Y = \sqrt{\hat{\sigma}_Y^2} = \sqrt{1/T \sum_{t=1}^{T} (Y_i - \overline{Y})^2)}$, hence $\hat{\sigma}_Y^2$ is the variance of Y (and similarly for $\hat{\sigma}_X$).

Exercise 1.7. As a first example of real-world time series, we will look at data collected from the Fulton Fish market in New York (see Graddy, 1995; 2006). The dataset we use involves aggregated daily prices and quantities of whiting for the period 2 December 1991 to 8 May 1992.[4]

(1) Download the file Fulton.zip provided on the book's webpage to get the access to the dataset fish.xls (or one of the other data bank formats provided). The variable definitions to note are as follows:

- *pricelevel*: the average daily price of whiting in USD.
- *price*: the natural logarithm of *price level*.
- *tots*: total quantity sold (Daily aggregated quantity in pounds)
- *qty*: the natural logarithm of *tots*.
- *stormy*: A dummy (indicator variable) which is 1 if the weather at sea was stormy, and zero otherwise.
- *mixed*: A dummy (indicator variable) which is 1 if the weather at sea was mixed, and zero otherwise.

[4]The dataset was downloaded from http://people.brandeis.edu/~kgraddy/datasets/fish.out on 6 July 2018.

A dummy variable for public holiday has been added to the dataset following the analysis in Hendry and Nielsen (2007, p. 200)[5]:

- *hol*: 1 if trading day is on or close to a public holiday. Zero otherwise.

(2) Investigate the two time series *pricelevel* and *tots* for signs of autocorrelation.

(3) What about *price* and *qty*?

[5]Two other existing analyses of this dataset are Angrist *et al.* (2000) and Graddy and Kennedy (2010).