1

Introduction

This is a book on Growth Theory and on the numerical methods needed to fully characterize the properties of most Growth models. In this introductory chapter, we describe the main characteristics of different families of Growth models and their relevance for policy analysis, which is moving leading economic and financial institutions throughout the world to increasingly rely on their use for forecasting as well as for policy evaluation. In particular, we emphasize how the richer structure provided to Growth models by their Microeconomic foundations allows us to address a much broader set of policy issues than in more traditional structural dynamic models. The book gradually builds on by increasing the degree of generality of the models being considered, as explained below. We cover: i) neoclassical growth under a constant savings rate, ii) optimal growth, iii) numerical solution methods, iv) endogenous growth, and v) monetary growth. Theoretical discussions on each model are presented, with special attention to characterizing the properties of equilibrium solutions and their use for fiscal policy considerations, while a specific chapter deals with monetary policy issues. Algorithms to solve all models considered are presented, together with EXCEL spreadsheets and MATLAB programs that implement them. Results obtained by these programs are commented in '*Numerical exercise*'-type sections, where some indications are provided on possible modifications of the enclosed programs. The book has been written with the intention that it may be accessible to students without an initial background on Growth Theory or mathematical software. Maintaining the same notation used in the analytical presentations in the book should allow the reader to follow easily the structure of the programs and quickly learn how to adapt them to alternative specifications or theoretical assumptions.

Growth models incorporate very specific assumptions on the structure of preferences, technology, the sources of randomness, and the policy rules followed by the economic authority, and characterize the relationship implied by such a structure between the decisions made by the different agents at each point in time and the information they have available when making their decisions. Under uncertainty, agents' perceptions on the future are an explicit determinant of their actions. Growth models do not make adhoc assumptions on the way how expectations influence agents' decisions. Rather, the solution to the optimization problems posed for each agent leads to decision rules for the different agents that incorporate expectations of functions of future variables in a very specific manner. If expectations are assumed to be rational, expectations in the model become endogenous variables, they are fully consistent with the structure of the model, and incorporate agents' perceptions of possible future changes in policy. Doing that, these models are safe from a strong criticism made on a traditional approach to economic policy evaluation by Nobel laureate R.E. Lucas that has been very influential in the last decades. This is the reason why, as we describe below, these models are increasingly being used in the research departments of Central Banks and main international economic institutions to forecast as well as to evaluate the consequences of alternative policy choices.

The counterpart comes from the fact that the type of stochastic control problems that are integrated into a Growth model lack an analytical solution, so they need to be solved following a numerical approach, accompanied by Monte Carlo simulation in the case of stochastic Growth models. The numerical solution to the model then comes in the form of artificial time series that can be analyzed using standard statistical and econometric tools, and the results compared to those obtained in corresponding time series data from actual economies. These are the main issues introduced in this chapter, which are later gradually developed throughout the book. Section 1 reviews some statistical concepts using simple time series models, Section 2 considers some simple dynamic macroeconomic models in which we introduce additional concepts, as well as the fundamentals of the simulation methods that will be used through the book. Section 3 introduces the main characteristics of Growth models, in comparison with more traditional dynamic macroeconomic models. This section motivates the convenience to work with Growth models and describes their different types, paying attention to the way they deal with the criticism to more traditional policy evaluation. Section 4 explains the need to obtain numerical solutions to Growth models, their potential use, and how this approach has led to changing the type of policy questions we ask and the type of answers we get. This introductory chapter ends up with a synopsis of the book, where a reference is made to the treatment of the issues mentioned along this Introduction.

1.1 A few time series concepts

Economics is full of statements relating the dynamic properties of key variables. For instance, we may say that inflation is very persistent, that aggregate consumption and GNP experience cyclical fluctuations, or that hours worked and productivity move independently from each other. These statements have direct implications in terms of the time series representations of these variables. Sometimes we are more specific, as when we state that stock exchange returns are white noise, thereby justifying the usual belief that they are *unpredictable*. The unpredictability statement comes from the fact that the forecast of a white noise process, no matter how far into the future, is always the same. That forecast is equal to the mean of the white noise process, which would likely be assumed to be zero in the case of asset returns. If returns are logarithmic, i.e., the first difference of logged market prices, then prices themselves would follow a random walk structure. These properties cannot be argued separately from each other, since they are just two different forms of making the same statement on stock market prices. We may also say at some point that the economy is likely to repeat next year its growth performance from the previous year, which incorporates the belief that annual GNP growth follows a random walk, its best one-step ahead prediction being the last observed value. A high persistence in real wages or in inflation could be consistent with first order autoregressive models with an autoregressive parameter close to 1. We briefly review in this section some concepts regarding basic stochastic processes, of the type that are often used to represent the behavior of economic variables.

1.1.1 Some simple stochastic processes

A stochastic process is a sequence of random variables indexed by time. Each of the random variables in a stochastic process, corresponding to a given time index t, has its own probability distribution. These distributions can be different, and any two of the random variables in a stochastic process may either exhibit dependence of some type or be independent from each other.

A *white noise* process is,

$$y_t = \varepsilon_t, \quad t = 1, 2, 3, ...$$

where $\varepsilon_t, t = 1, 2, ...$ is a sequence of independent, identically distributed zero-mean random variables, known as the *innovation* to the process. A white noise is sometimes defined by adding the assumption that ε_t has a Normal distribution. The mathematical expectation of a white noise is zero, and its variance is constant: $Var(y_t) = \sigma_\varepsilon^2$. More generally, we could consider a *white noise with drift*, by incorporating a constant term in the process,

$$y_t = a + \varepsilon_t, \quad t = 1, 2, 3, ...$$

with mathematical expectation $E(y_t) = a$, and variance: $Var(y_t) = \sigma_\varepsilon^2$.

The future value of a white noise with drift obeys,

$$y_{t+s} = a + \varepsilon_{t+s},$$

so that, if we try to forecast any future value of a white noise on the basis of the information available[1] at time t, we would have:

$$E_t y_{t+s} = a + E_t \varepsilon_{t+s} = a,$$

[1] That amounts to constructing the forecast by application of the conditional expecta-

because of the properties of the ε_t-process. That is, the prediction of a future value of a white noise is given by the mean of the process. In that sense, a white noise process is *unpredictable.* The prediction of such process is given by the mean of the process, with no effect from previously observed values. Because of that, the history of a white noise process is irrelevant to forecast its future values. No matter how many data points we have, we will not use them to forecast a white noise.

A *random walk with drift* is a process,

$$y_t = a + y_{t-1} + \varepsilon_t, \quad t = 1, 2, 3, ... \tag{1.1}$$

so that its first differences are white noise. If $y_t = \ln(P_t)$ is the log of some market price, then its return $r_t = \ln(P_t) - \ln(P_{t-1})$, will be a white noise, as we already mentioned. A random walk does not have a well defined mean or variance.

In the case of a *random walk without drift*, we have,

$$y_{t+s} = y_{t+s-1} + \varepsilon_{t+s}, \quad s \geq 1$$

so that we have the sequence of forecasts:

$$\begin{aligned} E_t y_{t+1} &= E_t y_t + E_t \varepsilon_{t+1} = y_t, \\ E_t y_{t+2} &= E_t y_{t+1} + E_t \varepsilon_{t+2} = E_t y_{t+1} = y_t, \end{aligned}$$

and the same for all future variables. In this case, the history of a random walk process is relevant to forecast its future values, but only through the last observation. All data points other than the last one are ignored when forecasting a random walk process.

First order autoregressive processes, AR(1), are of the form,

$$y_t = \rho y_{t-1} + \varepsilon_t, \ \mid \rho \mid < 1,$$

and can be represented by,

$$y_t = \sum_{s=0}^{\infty} \rho^s \varepsilon_{t-s},$$

the right hand side having a finite variance under the assumption that $Var(\varepsilon_t) = \sigma_\varepsilon^2$ only if $|\rho| < 1$. In that case, we would have:

$$E(y_t) = 0; \quad Var(y_t) = \frac{\sigma_\varepsilon^2}{1 - \rho^2}.$$

---

tion operator to the analytical representation of the future value being predicted, where the conditional expectation is formed with respect to the sigma algebra of events known at time t.

Predictions from a first order autoregression can be obtained by,

$$\begin{aligned} E_t y_{t+1} &= \rho E_t y_t + E_t \varepsilon_{t+1} = \rho y_t, \\ E_t y_{t+2} &= E_t(\rho y_{t+1}) + E_t \varepsilon_{t+2} = \rho^2 E_t y_{t+1} = \rho^2 y_t, \end{aligned}$$

and, in general,

$$E_t y_{t+s} = \rho^s y_t, \quad s \geq 1$$

which is the reason to impose the constraint $\mid \rho \mid < 1$. The parameter ρ is sometimes known as the *persistence* of the process. As the previous expression shows, an increase or decrease in y_t will show up in any future y_{t+s}, although the influence of that y_t-value will gradually disappear over time, according to the value of ρ. A value of ρ close to 1 will therefore introduce high persistence in the process, the opposite being true for ρ close to zero.

The covariance between the values of the first order autoregressive process at two points in time is:

$$Cov(y_t, y_{t+s}) = \rho^s Var(y_t), \; s \gtreqless 0,$$

so that the linear correlation is:

$$Corr(y_t, y_{t+s}) = \frac{Cov(y_t, y_{t+s})}{Var(y_t)} = \rho^s,$$

which dies away at a rate of ρ. In an autoregressive process with a value of ρ close to 1, the correlation of y_t with past values will be sizeable for a number of periods.

A first order autoregressive process with constant has the representation,

$$y_t = a + \rho y_{t-1} + \varepsilon_t, \; \mid \rho \mid < 1.$$

Let us assume by now that the mathematical expectation exists and is finite. Under that assumption, $Ey_t = Ey_{t-1}$, and we have:

$$Ey_t = a + E(\rho y_{t-1}) + E\varepsilon_t = a + \rho E y_t,$$

so that: $Ey_t = \frac{a}{1-\rho}$. To find out the variance of the process, we can iterate on its representation:

$$\begin{aligned} y_t &= a + \rho y_{t-1} + \varepsilon_t = a + \rho(a + \rho y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\ & a(1 + \rho + \rho^2 + ... + \rho^{s-1}) + \rho^s y_{t-s} \\ & + \left(\rho^{s-1}\varepsilon_{t-s+1} + ... + \rho^2 \varepsilon_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t\right), \end{aligned}$$

and if we proceed indefinitely, we get

$$y_t = a(1 + \rho + \rho^2 + ...) + \left(... + \rho^2 \varepsilon_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t\right),$$

since $\lim_{s\to\infty} \rho^s y_{t-s} = 0$.[2] Then, taking the variance of this expression:

$$Var(y_t) = Var\left(... + \rho^2\varepsilon_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t\right) = \sum_{s=0}^{\infty} \rho^{2s}\sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2},$$

so that the variance of the y_t-process increases with the variance of the innovation, σ_ε^2, but it is also higher the closer is ρ to 1. As ρ approaches 1, the first order autoregression becomes a random walk, for which this expression would give an infinite variance. This is because if we repeat for the random walk the same argument we have made here, we get,

$$\begin{aligned} y_t &= a + y_{t-1} + \varepsilon_t = a + (a + y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\ & as + y_{t-s} + (\varepsilon_{t-s+1} + ... + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t), \end{aligned}$$

so that the past term y_{t-s} does not die away no matter how far we move back into the past, and the variance of the sum in brackets increases without bound as we move backwards in time. The random walk process has an infinite variance. Sometimes, it can be assumed that there is a known initial condition y_0. The random walk process can then be represented:

$$\begin{aligned} y_t &= a + y_{t-1} + \varepsilon_t = a + (a + y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\ &= ... = at + y_0 + (\varepsilon_1 + ... + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t), \end{aligned}$$

with $E(y_t) = ta$ and $Var(y_t) = t\sigma_\varepsilon^2$. Hence, both moments change over time, the variance increasing without any bound. However, if we compare in a same graph time series realizations of a random walk together with some stationary autoregressive processes, it will be hard to tell which is the process with an infinite variance.

A future value of the first order autoregression can be represented:

$$y_{t+s} = a + \rho y_{t+s-1} + \varepsilon_{t+s}, \ \mid \rho \mid < 1, \ \ s \geq 1,$$

which can be iterated to,

$$y_{t+s} = a(1+\rho+\rho^2+...+\rho^{s-1}) + \rho^s y_t + \left(\rho^{s-1}\varepsilon_{t+1} + \rho^{s-2}\varepsilon_{t+2} + ... + \varepsilon_{t+s}\right),$$

so that its forecast is given by,

$$y_{t+s} = a\frac{1-\rho^s}{1-\rho} + \rho^s y_t.$$

So, as the forecast horizon goes to infinity, the forecast converges to,

$$\lim E_t y_{t+s} = \frac{a}{1-\rho},$$

the mean of the process.

[2] This is the limit of a random variable, and an appropriate limit concept must be used. It suffices to say that the power of ρ going to zero justifies the zero limit for the product random variable.

1.1.2 Stationarity, mean reversion, impulse responses

A stochastic process is stationary when the distribution of k-tuples $(y_{t_1}, y_{t_2}, ..., y_{t_k})$ is the same with independence of the value of k and of the time periods $t_1, t_2, ..., t_k$ considered. It is a property of any *stationary stochastic process* that the forecast of a future value converges to its mean as the forecast horizon goes to infinity. This is obviously fulfilled in the case of a white noise process. Another characteristic is that any time realization crosses the sample mean often, while a nonstationary process would spend arbitrarily large periods of time at either side of its sample mean. As we have seen above for the first order autoregression, the simple autocorrelation function of a stationary process, made up by the sequence of correlations between any two values of the process, will go to zero relatively quickly, dieing away very slowly for processes close to nonstationarity.

When they are not subject to an stochastic innovation,[3] stationary autoregressive processes converge smoothly and relatively quickly to their mathematical expectation. The y_t-process will converge to $\frac{a}{1-\rho}$ either from above or from below, depending on whether the initial value, y_0, is above or below $\frac{a}{1-\rho}$. The speed of convergence is given by the autoregessive coefficient. When the process is subject to a nontrivial innovation, the convergence in the mean of the process will not be easily observed. This is the case because the process experiences a shock through the innovation process every period, which would start a new convergence that would overlap the previous one, and so on. Under normal circumstances we will just see a time realization exhibiting fluctuations around the mathematical expectation of the process, unless the process experiences a huge innovation, or the starting condition y_0 is far enough from $\frac{a}{1-\rho}$, in units of its standard deviation, $\sqrt{\frac{\sigma_\varepsilon^2}{1-\rho^2}}$.

The property of converging to the mean after any stochastic shock is called *mean reversion*, and is characteristic of stationary processes. In stationary processes, any shock tends to be corrected over time. This cannot be appreciated because shocks to y_t are just the values of the innovation process, which take place every period. So, the process of mean reversion following a shock gets disturbed by the next shock, and so on. But the stationary process will always react to shocks as trying to return to its mean. Alternatively, a non stationary process will tend to depart from its mean following any shock. As a consequence, the successive values of the innovation process ε_t will take y_t every time farther away from its mean.

An alternative way of expressing this property is through the effects of purely transitory shocks or innovations. A stationary process has transitory responses to purely transitory innovations. On the contrary, a nonstationary process may have permanent responses to purely transitory shocks. So,

[3] That is, if the inovation ε_t has zero variance.

if a stationary variable experiences a one-period shock, its effects may be felt longer than that, but will disappear after a few periods. The effects of such a one-period shock on a nonstationary process will be permanent. A white noise is just an innovation process. The value taken by the white noise process is the same as that taken by its innovation. Hence, the effects of any innovation last as long as the innovation itself, reflecting the stationary of this process. The situation with a random walk is quite different. A random walk takes a value equal to the one taken the previous period, plus the innovation. Hence, any value of the innovation process gets accumulated in successive values of the random walk. The effects of any shock last forever, reflecting the nonstationary nature of this process. In a stationary first order autoregression, any value of the innovation ε_t gets incorporated into y_t that same period. It will also have an effect of size $\rho\varepsilon_t$ on y_{t+1}. This is because $y_{t+1} = \rho y_t + \varepsilon_{t+1}$ so, even if $\varepsilon_{t+1} = 0$, the effect of ε_t would still be felt on y_{t+1} through the effect it previously had on y_t.

This argument suggests how to construct what we know as an *impulse response function*. In the case of a single variables, as with the stochastic processes we consider in this section, that response is obtained by setting the innovation to zero every period except by one, in which the impulse is produced. At that time, the innovation takes a unit value.[4] The impulse response function will be the difference between the values taken by the process after the impulse in its innovation, and those that would have prevailed without the impulse. The response of a white noise to an impulse in its own innovation is a single unit peak at the time of the impulse, since the white noise is every period equal to its innovation, which is zero except at that time period. In the case of a general random walk, a zero innovation would lead to a random walk growing constantly at a rate defined by the drift a from a given initial condition y_0. If at time t^* the innovation takes a unit value, the random walk will increase by that amount at time t^*, but also at any future time. So the impulse response is in this case a *step function*, that takes the value 1 at t^* and at any time after that. Consider now a stationary first order autoregression. A unit innovation at time t^* will have a unit response at that time period, and a response of size ρ^s each period $t + s$, gradually decreasing to zero.

Another important characteristic of economic time series is the possibility that they exhibit cyclical fluctuations. In fact, first order autoregressive processes may display a shape similar to that of many economic time series, although to produce regular cycles we need a second order autoregressive processes,

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t,$$

[4] When working with several variables, responses can be obtained for impulses in more than one variable. To make the size of the responses comparable, each innovation is supposed to take a value equal to its standard deviation, which may be quite different for different innovations.

with ε_t being an innovation, a sequence of independent and identically distributed over time. Using the lag operator: $B^s y_t = y_{t-s}$ in the representation of the process:

$$y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} = \left(1 - \rho_1 B - \rho_2 B^2\right) y_t = \varepsilon_t.$$

The dynamics of this process is characterized by the roots of its characteristic equation,

$$1 - \rho_1 B - \rho_2 B^2 = (1 - \lambda_+ B)(1 - \lambda_- B) = 0,$$

which are given by:

$$\lambda_+, \lambda_- = \frac{-\rho_1 \pm \sqrt{\rho_1^2 + 4\rho_2}}{2\rho_2}.$$

Stationary second order autoregressions have the two roots of the characteristic equation smaller than 1. A root greater than one in absolute size will produce an explosive behavior. A root equal to one also signals nonstationarity, although the sample realization will not be explosive. It will display extremely persistent fluctuations, very rarely crossing its mean, as it was the case with a random walk. This is very clear in the similar representation of a random walk: $(1 - B)\, y_t = \varepsilon_t$.

Since the characteristic equation is now of second degree, it might have as roots two conjugate complex numbers. When that is the case, the autoregressive process displays cyclical fluctuations. The response of y_t to an innovation ε_t will also display cyclical fluctuations, as we will see in dynamic macroeconomic models below.

1.1.3 Numerical exercise: Simulating simple stochastic processes

The *Simple_simul.xls* EXCEL book presents simulations of some of these simple stochastic processes. Column A in the Simulations spreadsheet contains a time index. Column B contains a sample realization of random numbers extracted from a $N(0,1)$ distribution. This has been obtained from EXCEL using the sequence of keys: *Tools/Data Analysis/Random Number Generator* and selecting as options in the menu *number of variables* =1, *observations* = 200, a *Normal* distribution with expectation 0 and variance 1, and selecting the appropriate output range in the spreadsheet.

A well constructed random number generator produces independent realizations of the chosen distribution. We should therefore have in column B 200 independent data points from a N(0,1), which can either be interpreted as a sample of size 200 from a N(0,1) population, or as a single time series realization from a white noise where the innovation follows a N(0,1) probability distribution. The latter is the interpretation we will follow. At the end of the column, we compute the sample mean and standard deviation, with values of 0.07 and 1.04, respectively. These are estimates of the 0 mathematical expectation and unit standard deviation with this sample. Below that, we present the standard deviation of the first and the last 100 observations, of 1.09 and .98. Estimates of the variance obtained with the full sample or with the two subsamples seem reasonable. A different sample would lead to different numerical estimates.

Panel 2 contains sample realizations from three different random walks without drift. The only parameter in such processes is the variance of the innovation, which takes values 1, 25 and 100, respectively. At a difference of a white noise, an initial condition is needed to generate a time series for a random walk, because of the time dependence between successive observations, as can be seen in (1.1) . The three sample realizations are graphed in the *RandomWalks* spreadsheet. All exhibit extreme persistence, crossing the sample mean just once in 200 observations. We know by construction that these three processes lack a well defined mean and have a time increasing variance. We can always compute sample averages and standard deviations, as shown in the spreadsheet at the end of the series, but it is not advisable to try to interpret such statistics. In particular, in this case, by drawing different realization for the white noise in column B, the reader can easily check how sample mean and standard deviations may drastically change. In fact, standard deviations are calculated in the spreadsheet for the first and last 100 sample observations, and they can turn out to be very different, and different from the $t\sigma_\varepsilon^2$ theoretical result. The point is we cannot estimate that time-varying moment with much precision.

Panel 3 compares a random walk to three first-order autoregressive processes, with autoregressive coefficients of 0.99, 0.95 and 0.30. As mentioned above, a random walk can be seen as the limit of a first order autoregression, as the autoregressive coefficient converges to 1, although the limit presents some discontinuity since, theoretically, autoregressive processes are stationary so long as the autoregressive coefficient is below 1 in absolute value, while the random walk is nonstationary. The autoregressive processes will all have a well-defined mean and variance, which is not the case for the limit random walk process. 0.99. The sample time series realizations for the four processes are displayed in the *AR-processes* spreadsheet, where it can be seen that sample differences between the autoregressive process with the 0.99 coefficient and the random walk are minor, in spite of the theoretical differences between the two processes. In particular, the autoregressive process crosses its sample mean in very few occasions. That is also the case for the 0.95-autoregressive process, although its mean reverting behavior is very clear at the end of the sample. On the other hand, the time series realization from the 0.30-autoregressive process exhibits the typical behavior in a clearly stationary process, crossing its sample mean repeatedly.

Panel 4 presents sample realizations from two white noise processes with drift and N(0,1) innovations. As shown in the enclosed graph, both fluctuate around their mathematical expectation, which is the value of the constant defining the drift, crossing their sample means very often. Panel 5 contains time series realizations for two random walk processes with drift. These show in the graph in the form of what could look as deterministic trends. This is because the value of the drifts, of 1.0 and 3.0, respectively, is large, relative to the innovation variance which is of 25 in both cases. If the value of the drift is reduced, or the variance of the innovation increased, the shape of the time series would be different, since the fluctuations would then dominate over the accumulated effect of the drift, as the reader can check by reducing the numerical values of the drift parameters[5] used in the computation of these two columns.

Panel 6 presents realizations of a stationary first order autoregression with coefficient of .90. In the second case we have not included an innovation process, so that it can be considered as a deterministic autoregression. It is interesting to see in the enclosed graph the behavior of a stationary process: starting from an initial condition, in the absence of an innovation, the process will always converge smoothly to its mathematical expectation. That is not the case in the stochastic autoregression, just because the innovation variance, of 25, is large relative to the distance between the initial condition, 150, and the mathematical expectation, 100. The reader can check how reducing the standard deviation used in column S from 5 to 0.5, the pattern of the time series changes drastically, and the convergence process becomes then evident.

Panel 7 contains realizations for second order autoregressions. The first two columns present sample realizations from stationary autoregressions,

$$\text{Model 1: } y_t = 10 + .6y_{t-1} + .3y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1.2}$$

$$\text{Model 2: } y_t = 30 + 1.2y_{t-1} - .5y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1.3}$$

and are represented in an enclosed graph. The two time series display fluctuations around their sample mean of 100, which they cross a number of times. The second time series, represented in red in the graph can be seen to exhibit a more evident stationary behavior, with more frequent crosses with the mean. The next three columns present realizations for nonstationary second order autoregressions. There is an important difference between them: the first two correspond to processes:

$$\text{Model 3:} \quad y_t = .7y_{t-1} + .3y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1.4}$$

$$\text{Model 4:} \quad y_t = 1.5y_{t-1} - .5y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1.5}$$

that contain exactly a unit root, the second one being stable.[6] The rots

[5] Or significantly increasing the innovation variance. What are the differences beetwen both cases in terms of the values taken by the process?

[6] The two polynomials can be written as $1 - a_1B - a_2B^2 = (1-B)(1-\lambda B)$, the second root being $1/\lambda$. The reader just need to find the value of λ in each case.

of the characteristic equation for Model 3 are 1 and -0.3, while those for Model 2 are 1 and 0.5. The last autoregression

$$\text{Model 5} : y_t = .3y_{t-1} + 1.2y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1.6}$$

has a root greater than one, which produces an explosive behavior. The two roots are -0.95 and 1.25.

The *Impulse responses* spreadsheet contains the responses to a unit shock for the stochastic processes considered above: a random walk, three first-order autoregressions, two stationary second-order autoregressions, and three nonstationary second-order autoregressions. The innovation in each process is supposed to take a zero value in each case for ten periods, to be equal to 1, the standard deviation assumed for the innovation in all cases at $t^* = 11$, and be again equal to zero afterwards. We compare that to the case when the innovation is zero at all time periods. Impulse responses are computed as the difference between the time paths followed by each process under the scenario with a shock at $t^* = 11$, and in the absence of that shock. The first-order autoregressions are supposed to start from an initial condition $y_0 = 100$, when their mathematical expectations is zero, so in the absence of any shock, they follow a smooth trajectory gradually converging to zero at a speed determined by its autoregressive coefficient. The second order autoregressions are assumed to start from $y_0 = y_1 = 100$, which is also their mathematical expectations. So, in the absence of any shock, the processes would stay at that value forever.[7]

The first graph to the right displays impulse responses for a random walk as well as for the three first order autoregressions considered above, with coefficients 0.99, 0.95 and 0.30. A random walk has the constant, permanent impulse response that we mentioned above when describing this process. The responses of the first order autoregressions can be seen to gradually decrease to zero from the initial unit value. The response is shorter the lower it is the autoregressive coefficient. For high autoregressive coefficients, the process shows strong persistence, which makes the effects of the shock to last longer.

The second graph shows the impulse responses of the two stationary second-order autoregressions. As the reader can easily check, the characteristic equation for Model 1 has roots -0.32 and 0.92, so it is relatively close to nonstationarity. The characteristic equation for Model 2 has roots $0.6 \pm 0.374\,17i$, with modulus 0.5. This difference shows up in a much more persistent response of Model 1. The complex roots of Model 2 explain the oscillatory behavior of the impulse response of this model.

The third graph displays impulse responses for the three nonstationary second order autoregressions. In the two cases when there is a unit root

[7] We could have done otherwise, like starting the first-order autoregresisons at their mathematical expectation, and the second-order autoreegressions outside their expected values. The reader can experiment with these changes.

(Models 3 and 4), the graph shows a permanent response to the purely transitory, one-period shock. The response of Model 5 is explosive because of having one root above 1, and its values are shown on the right Y-axis.

1.2 Structural macroeconomic models

In this section we review the main characteristics of structural macroeconomic models, paying special attention to some of the statistics summarizing their properties, since they will also be used to analyze Growth models. Structural models are specified as a system of relationships that include decision rules by economic agents, policy rules, and identities. The first ones are supposed to have originated in an optimizing behavior on the part of economic agents, which is never made explicit. We will focus or attention to dynamic structural models although, to have an appropriate perspective, we nevertheless start with a reference to static macroeconomic models.

1.2.1 Static structural models

A linear, static model is made up by a set of equations in which all variables are supposed to refer to the same time period, so that there is no need to use time indexes. Nevertheless, the model is interpreted as relating the values taken by endogenous and exogenous variables at each point in time. A solution to the model is a representation of endogenous variables as functions of structural parameters and exogenous variables only. When such a representation exists, the model can be used to actually compute implied values for endogenous variables as a function of given values for exogenous variables and parameters. A necessary condition for a linear, static model to have a solution is that it must have as many equations as endogenous variables. An example of such a model, in logged variables, is:

$$
\begin{aligned}
n &= \frac{d_0 + a_2 \bar{k} - (w - p)}{1 - a_1}, \\
n &= \eta \left(w - p \right), \\
y &= a_0 + a_1 n + a_2 \bar{k}, \\
y &= \left[c_1 (1 - \tau) y - c_2 (r - \pi^e) \right] + \left[i_1 - i_2 (r - \pi^e) \right] + \bar{g}, \\
\bar{m} - p &= m_1 y - m_2 r.
\end{aligned}
$$

The equations in this system are: i) the demand for labor,[8] increasing in the stock of capital and decreasing in the real wage, ii) the supply of labor,

[8]As it would be obtained by a profit-maximizing competitive firm with a Cobb-Douglas technology, $Y = a_0 K^{a_1} L^{a_2}, a_1 + a_2 \leq 1$, represented in logs by the first relationship, with $d_0 = \ln(a_0 a_1)$.

increasing in the real wage, iii) the production function, that determines the supply of goods, iv) the aggregate demand for goods, made up by the private demand for consumption and investment (both inversely related to the real rate of interest), plus government expenditures, which are assumed to be given at $\bar{g}$, and v) the market clearing condition in the money market, where the supply of real balances is $\bar{m}-p$, with $\bar{m}$ fixed by monetary policy. Market clearing conditions for the labour and goods markets have already been imposed by using the same notation for demand and supply variables. Endogenous variables are $n, y, w-p, p, r$, while exogenous variables are the stock of capital $\bar{k}$, expected inflation, π^e, money supply, $\bar{m}$, and government expenditures, $\bar{g}$. The income tax rate, τ, is one of the parameters of the model, together with input shares in production, or the elasticities in the money demand function.

This model has a recursive structure that allows for a simple analytical solution. The first two equations, labour demand and supply equations, determine the levels of employment and the real wage, the third equation determines the level of output, the equilibrium condition in the goods market determines interest rates, and the equilibrium condition in the money market determines the price level. The solution is:

$$\begin{aligned}
w-p &= \omega_0+\omega_1\bar{k}; \quad n=\eta\omega_0+\eta\omega_1\bar{k};\\
\omega_0 &= \frac{d_0}{1+\eta(1-a_1)}; \quad \omega_1=\frac{a_2}{1+\eta(1-a_1)};\\
y &= Y_0+K_0\bar{k}; \quad Y_0=a_0+\frac{a_1 d_0\eta}{1+\eta(1-a_1)}; \quad K_0=\frac{a_2(1+\eta)}{a_0+\frac{a_1 d_0 \eta}{1+\eta(1-a_1)}};\\
r &= \pi^e+\frac{i_1+\bar{g}}{c_2+i_2}-R_0Y_0-R_0K_0\bar{k}; \quad R_0=\frac{1-c_1(1-\tau)}{c_2+i_2};\\
p &= \bar{m}+m_2\pi^e-(m_1+m_2R_0)K_0\bar{k}-(m_1+m_2R_0)Y_0+m_2\frac{i_1+\bar{g}}{c_2+i_2}.
\end{aligned}$$

It is immediate to see that an increase of a unit in government expenditures would raise nominal and real interest rates by $\frac{1}{c_2+i_2}$, and the price level by $\frac{m_2}{c_2+i_2}$, with no effect on employment or output. An increase in money supply would raise the price level in the same amount, without affecting any other variable, showing the neutrality of money in this model. Alternative policy exercise could be conducted on the solution without any difficulty, the same way we could explore the potential effects of changes in the elasticity of an input in the aggregate production function, or changes in any elasticity in the consumption, investment or money demand functions. There are two ways to work with this model: i) the way it is specified, it is better conceived as a long-run model, that is solved under alternative values of exogenous variables and parameters to obtain long-run equilibria values for endogenous variables. When values for endogenous variables are calculated again after introducing some changes in exogenous variables or parameters, we would interpret the result as the equilibrium that would prevail in the economy after those changes have been implemented and enough time has passed for the equilibrium to be restored. From this point of view, the model is silent with respect to short-run adjustments. An alternative use of the model would assume time paths for exogenous variables $\bar{k}$, π^e, $\bar{m}$, $\bar{g}$, and values for structural parameters like the income tax rate, τ,to compute implied time paths for the vector of endogenous variables, $n, y, w-p, p, r$. That way, the implications of this static model could be compared with some statistical properties observed in time series data. In this particular model, a constant stock of capital is a short-run type of assumption, that suggests a preference for the first interpretation. If the model is to be used to relate variables over a long time span, an investment equation should better be added.

In general, a linear static model can be written: $Ay = B + Cx$, where x is the kx1 vector of exogenous variables, and y is the $n \times 1$ vector of endogenous variables, A is $n \times n$, B is $n \times 1$, and C is $n \times k$. in the previous example: $y = (n, y, w-p, p, r)'$, $x = (\bar{k}, \pi^e, \bar{m}, \bar{g})'$, and

$$A = \begin{pmatrix} 1-a_1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -\eta & 0 & 0 \\ -a_1 & 1 & 0 & 0 & 0 \\ 0 & 1-c_1(1-\tau) & 0 & 0 & c_2+i_2 \\ 0 & -m_1 & 0 & -1 & m_2 \end{pmatrix};$$

$$B = \begin{bmatrix} d_0 \\ 0 \\ a_0 \\ i_1 \\ 0 \end{bmatrix}; \qquad C = \begin{pmatrix} a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & c_2+i_2 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Whenever matrix A has full rank, the model has as solution:

$$y = M + Nx, \text{ with } M = A^{-1}B, \; N = A^{-1}C. \tag{1.7}$$

Characterizing the solution to a nonlinear static model will usually be much harder. Such model takes the general form: $F(y_t, x_t; \theta) = 0$, with θ representing the vector of parameters, for which a representation like (1.7) will generally not exist. At each point in time, a numerical algorithm to solve nonlinear systems of equations should then be used to obtain the values of endogenous variables as a function of the values of exogenous variables and structural parameters. But a complete nonlinear system[9] of equations may have no solution, or have multiple solutions. In many cases, providing an answer to the question of interest in such a model would require computing a linear, log-linear or polynomial approximation to the

[9] A system with as many equations as endogenous variables.

$F(y_t, x_t) = 0$ system. The linear model above can be thought of as having this origin.

Stochastic models add random shocks to some equations, taking the form:[10]

$$Ay = B + Cx + D\varepsilon,$$

where ε is the rx1 vector of exogenous shocks, and D is nxr. If A has full rank, the model has as solution:

$$y = M + Nx + P\varepsilon, \text{ with } M = A^{-1}B,\ N = A^{-1}C,\ P = A^{-1}D. \tag{1.8}$$

When such a model admits a short-run interpretation, time series can be computed for endogenous variables, contingent on a given scenario for the future evolution of exogenous variables and on some sample realizations for the exogenous shocks, given some values for structural parameters. Sample realizations for the exogenous shocks will be obtained by Monte Carlo simulation, under some assumption on their probability distribution, as it is explained below. Then, the model relates mean values of endogenous and exogenous variables, and the variance of endogenous variables to the variance of exogenous variables and innovations. The model will also have implications regarding the linear correlation coefficients between pairs of variables.[11] The number of innovations in the model, r, will limit the dimensionality of a statistical system that can be analyzed with the variables of the model. For instance, if $r = 1$, then any system with two or more equations, estimated with the time series for exogenous and endogenous variables obtained form the solution procedure outlined above, would have a singular variance-covariance matrix for the random error terms. Specifications of this type have been used to analyze policy design under uncertainty, as in Poole (1970), who determined that nominal interest rates should be the preferred policy instrument when monetary or financial shocks (i.e., shocks to the LM-equation) are dominant, money supply being the best control policy when shocks on private or public consumption and investment shocks prevail (i.e,, shocks to the IS-equation).

1.2.2 Dynamic structural models

A dynamic macroeconomic model specifies endogenous variables as functions of *predetermined* variables (lagged endogenous variables), exogenous variables and exogenous shocks:

$$Ay_t = B + Cy_{t-1} + Dx_t + E\varepsilon_t,$$

[10] We assume here, for simplicity, that all random shocks are white noise. Extending the model to incorporate possible autoregressive structures for the shocks is straightforward.

[11] If we denote by p_i the i-th row of the nxr matrix P, then $Var(y_i) = p_i'\Sigma_\varepsilon p_i$, $Var(y_j) = p_j'\Sigma_\varepsilon p_j$, $Cov(y_i, y_j) = p_i'\Sigma_\varepsilon p_j$, and $Corr(y_i, y_j) = \frac{p_i'\Sigma_\varepsilon p_j}{\sqrt{p_i'\Sigma_\varepsilon p_i}\sqrt{p_j'\Sigma_\varepsilon p_j}}$, with Σ_ε being the rxr variance-covariance matrix of vector ε.

where variables have the same interpretation as above, except for the $n \times n$ matrix C of coefficients in predetermined variables. This first-order vector autoregressive representation can always be achieved by an appropriate definition of variables.[12] The *short-term solution* to the model would represent current endogenous variables as a function of exogenous variables, predetermined variables and structural parameters, and it would be obtained similarly to the static model, provided matrix A is invertible:

$$y_t = M + Ny_{t-1} + Px_t + Q\varepsilon_t,$$

with $M = A^{-1}B, \quad N = A^{-1}C, \quad P = A^{-1}D, \quad Q = A^{-1}E.$

As a static model, it can be simulated over time for specific trajectories of the exogenous variables, starting from initial conditions for predetermined variables. At a difference from static models, a dynamic macroeconomic model is intended to capture short-run fluctuations in endogenous variables, so that it has long- and short- term implications. The dynamics introduced by the presence of lagged endogenous variables implies that any policy intervention or structural change generally has nontrivial effects over some time period. Hence, these models have richer implications than purely static models, in the form of statistics like: short- and long-run multipliers, cross-correlations or impulse response functions, among others, not unlike those we have already seen in the statistical review of time series in the previous section.

The appropriate concept to analyze the implied *long-run relationships* between the values of endogenous and exogenous variables is that of steady-state, which we introduced below. A steady-state is obtained by setting $y_t = y_{t-1} = y^*$ while setting exogenous shocks to zero $\forall t$, and assuming constant exogenous variables at x^*, and solving the model for y^* as a function of x^*. Steady-state relationships from dynamic models are comparable to static models, which justifies their usual long-run interpretation. When long-run effects are the focus of interest, we just need to compare steady-states before and after a given structural change or policy intervention, that is, for alternative values of structural parameters or exogenous variables. While a static model can also establish that comparison, a dynamic model can describe the *transition*, i.e., the trajectory followed by endogenous variables between the old and the new steady-state. A dynamic model can be used to characterize not the duration of the transition, but also some major characteristics, like the time evolution of the rate of growth of output, interest rates or productivity along the transition. By describing the whole transition, dynamic macroeconomic models allow us to evaluate not only the long-term effects of structural changes and policy interventions, but also the effects along the transition. The policy maker will usually want

[12] If, for instance, C_t, C_{t-1} and C_{t-2} appear in the model, both, C_t and C_{t-1} will form part of vector y_t, while C_{t-1} and C_{t-2} will be included in vector y_{t-1}. The representation could also be extended easily to accomodate lagged innovation values.

to take into account the short- and the long-term consequences of any policy intervention. What makes this important is the fact that, as we will repeatedly see throughout this book, it is usually the case in dynamic models that a given policy intervention has effects of different sign on the short- than on the long-term, and either one can prevail, depending on the length of the transition, the size of both types of effects, and the rate of time discount. Hence, focusing on long-term effects alone, as it is done in static models, can easily provide a misleading answer to the policy analysis.

As an example, let us consider the model,

$$\begin{aligned} C_t &= \alpha_1 + \alpha_2 Y_{t-1}, \\ I_t &= \beta_1 + \beta_2 \left(Y_{t-1} - Y_{t-2} \right), \\ Y_t &= C_t + I_t + G_t, \end{aligned}$$

where C_t, I_t, Y_t, G_t denote private consumption and investment, output and government expenditures, respectively. The model has three equations and can therefore be used to explain the behavior of three endogenous variables. It seems natural that these should be consumption, investment and output. Moreover, the first equation can be labelled the consumption equation, explaining consumption as a function of last period's output/income. The second equation is investment, and can be interpreted as determining investment as a function of last period's changes in output, maybe because of adjustment costs of capital. The last equation is the national identity equation in a simple closed economy. This model is known in macroeconomics textbooks as a *multiplier-accelerator* model, since the second (investment) equation captures an acceleration effect in output. The two lags of output in the consumption and investment equations are *predetermined* as of time t, while public expenditures are considered to be *exogenous* to the model.

If we have data for current and future government expenditures, G_1, $G_2, G_3, ...$, as well as initial conditions on output Y_0, Y_{-1}, and parameter values α_1, α_2, β_1, β_2, the model contains enough information to provide us recursively with values for (C_1, I_1, Y_1), (C_2, I_2, Y_2), ... We would start obtaining C_1 from the consumption equation, I_1 from the investment equation, Y_1 from the national income identity, repeating the process for each time period. To do so, we will also need numerical values for the model's parameters, which may have been previously estimated using aggregate macroeconomic time series data. Alternatively, we could generate artificial time series data from the model following the procedure described, starting from some exogenously given initial conditions, and for hypothetical values of the structural parameters.

However, as it is well known, not any model is *identified*. To have the same number of equations as endogenous variables is a necessary, but not sufficient condition for the model to explain the behavior of the variables chosen as endogenous. To understand this, let us now suppose that we chose consumption, output and public expenditures as the endogenous variables. In that case, starting from known parameter values and a given path for investment I_1, I_2, $I_3, ...$ as well as initial values Y_0, Y_{-1}, we would again obtain C_1 from the consumption equation, but we would be left with the last equation to obtain values for G_1, Y_1, which is clearly impossible, reflecting the fact that with this choice of endogenous variables, the model would not be identified.

Coming back to the initial choice of endogenous variables, the iterative process we described for that case amounts to substituting the consumption and investment equations into the national income identity, to have,

$$Y_t - (\alpha_2 + \beta_2) Y_{t-1} + \beta_2 Y_{t-2} = (\alpha_1 + \beta_1) + G_t, \tag{1.9}$$

a second-order difference equation giving the current value of output as a function of its two previous values, as well as the current value of government expenditures. As shown in the next section, the second order polynomial in the left hand side of this equation can display many different types of behavior.

Dynamic behavior of endogenous variables

Let us suppose that, starting from initial values for output Y_0, Y_{-1}, government expenditures were fixed at a given value G^*, $G_t = G^* \; \forall t$. Even then, output would not be constant, in general. In fact, we would have:

$$\begin{aligned} Y_1 &= (\alpha_2 + \beta_2) Y_0 - \beta_2 Y_{-1} + (\alpha_1 + \beta_1) + G^*, \\ Y_2 &= (\alpha_2 + \beta_2) Y_1 - \beta_2 Y_0 + (\alpha_1 + \beta_1) + G^*, \\ Y_3 &= .... \end{aligned}$$

and whether output converges or explodes, i.e., whether it is stable or unstable, and whether it displays oscillations or not, depends just on the values of α_2 and β_2. It is interesting to point out that there is an equilibrium value of output, defined precisely as that level of output such that if the economy started there, it would never move away from it. When it exists, that point is also called the *steady-state* of the system. This equilibrium level can in fact be easily obtained. To do so, we assume output to be constant over time in (1.9), to obtain,

$$Y^* = \frac{(\alpha_1 + \beta_1) + G^*}{1 - \alpha_2},$$

which can be seen to be directly related to the level chosen for government expenditures. Corresponding to these equilibrium values of government expenditures and output there would be associated equilibrium values for private consumption and investment: $C^* = \alpha_1 + \alpha_2 \frac{(\alpha_1+\beta_1)+G^*}{1-\alpha_2}$, $I^* = \beta_1$. An economy could stay at equilibrium values G^*, Y^*, C^*, I^* forever.

However, if the economy stays at its equilibrium values, but government expenditures experiences some deviation from its equilibrium value G^*, to a new value G^{**}, the economy would then depart from values Y^*, C^*, I^*. It is then interesting to discuss whether the economy would converge to its new equilibrium value $Y^{**} = \frac{(\alpha_1+\beta_1)+G^{**}}{1-\alpha_2}$ or diverge away from it. If the economy converges, it is interesting to know whether it would display oscillations, or it would move along a smooth convergent path.

More specifically, the roots of the characteristic equation are,

$$\lambda_+, \lambda_- = \frac{(\alpha_2 + \beta_2) \pm \sqrt{(\alpha_2 + \beta_2)^2 - 4\beta_2}}{2},$$

so that the general solution to the homogeneous equation,

$$Y_t - (\alpha_2 + \beta_2) Y_{t-1} + \beta_2 Y_{t-2} = 0,$$

is,

$$Y_t = A_1 \lambda_+^t + A_2 \lambda_-^t,$$

showing that if either λ_+ or λ_- were greater than 1 in absolute value, then output will explode. Other possibilities are: a) λ_+ and λ_- are real, and less than 1 in absolute value. Then output converges monotonically to its new equilibrium, b) λ_+ and λ_- are conjugate complex numbers, less than 1 in absolute value. Output then converges to its new equilibrium displaying damped oscillations, c) λ_+ and λ_- are conjugate complex numbers, greater than 1 in absolute value. Output then presents explosive oscillations..

In summary, the solution will be stable if λ_+ and λ_- have both modulus less than 1, while if either one has modulus greater than 1, the solution will be unstable. The characteristic roots are complex if $4\beta_2 > (\alpha_2 + \beta_2)^2$.

The model could have been solved for either one of the other two endogenous variables, consumption and investment. For instance, using the consumption function to eliminate income values from (1.9), we would obtain,

$$C_t - (\alpha_2 + \beta_2) C_{t-1} + \beta_2 C_{t-2} = (\alpha_1 + \alpha_2 \beta_1) + \alpha_2 G_{t-1},$$

with the same characteristic equation as in the case of output, so that consumption will have the same dynamic properties as output in the solution to the model. This is a consequence of consumption being determined by the level of lagged output alone.

Dynamic multipliers

In the response of an endogenous variable to a change in the value of an exogenous variable, we distinguish between the initial effect (the impact multiplier), the response over time (the dynamic multipliers), and the aggregate response over time (the total long-run multiplier). We must also distinguish between the response to a transitory change in an exogenous variable and the response to a permanent change. In the case of the multiplier-accelerator model, the second order difference output equation can be written,

$$Y_t = (\alpha_2 + \beta_2) Y_{t-1} - \beta_2 Y_{t-2} + (\alpha_1 + \beta_1) + G_t, \tag{1.10}$$

that in first differences becomes,

$$\Delta Y_t = (\alpha_2 + \beta_2) \Delta Y_{t-1} - \beta_2 \Delta Y_{t-2} + \Delta G_t, \tag{1.11}$$

as can be seen by subtracting the versions of equation (1.10) corresponding to time t and $t-1$.

This equation clearly shows that the impact multiplier of a change in government expenditures is equal to 1, since any change in G_t translates into a change in output with coefficient 1. Obtaining the dynamic multipliers can be done by numerical simulation. Their analytical computation, is somewhat burdensome, since we need to perform iterative substitutions. We would start by writing (1.11) at time $t+1$,

$$\Delta Y_{t+1} = (\alpha_2 + \beta_2) \Delta Y_t - \beta_2 \Delta Y_{t-1} + \Delta G_{t+1},$$

$$\Delta Y_{t+2} = (\alpha_2 + \beta_2) \Delta Y_{t+1} - \beta_2 \Delta Y_t + \Delta G_{t+2},$$

and substitute (1.11) to obtain,

$$\begin{aligned}
\Delta Y_{t+1} &= \left[(\alpha_2 + \beta_2)^2 - \beta_2\right] \Delta Y_{t-1} - \beta_2 (\alpha_2 + \beta_2) \Delta Y_{t-2} \\
&\quad + \left[\Delta G_{t+1} + (\alpha_2 + \beta_2) \Delta G_t\right], \\
\Delta Y_{t+2} &= (\alpha_2 + \beta_2) \left[(\alpha_2 + \beta_2)^2 - 2\beta_2\right] \Delta Y_{t-1} \\
&\quad - \beta_2 \left[(\alpha_2 + \beta_2)^2 + \beta_2\right] \Delta Y_{t-2} \\
&\quad + \left[\Delta G_{t+2} + (\alpha_2 + \beta_2) \Delta G_{t+1} + \left[(\alpha_2 + \beta_2)^2 - \beta_2\right] \Delta G_t\right],
\end{aligned}$$

where variations in output previous to time t are zero, $\Delta Y_{t-1} = \Delta Y_{t-2} = 0$.

We must distinguish two different cases:

i) If the change in government expenditures was *permanent*, and of size 2, we would have:

$$\Delta G_t = 2, \ \Delta G_{t+1} = \Delta G_{t+2} = ... = 0,$$

with an output response,

$$\begin{aligned}
\Delta Y_t &= 2, \ \Delta Y_{t+1} = 2 (\alpha_2 + \beta_2), \\
\Delta Y_{t+2} &= 2 \left[(\alpha_2 + \beta_2)^2 - \beta_2\right], \ ...
\end{aligned}$$

ii) On the other hand, if the change in government expenditures was purely *transitory*, lasting for just one period, and was of size 2, we will have,

$$\Delta G_t = 2,\ \Delta G_{t+1} = -2,\ \Delta G_{t+2} = ... = 0,$$

with an output response,

$$\Delta Y_t = 2,\ \Delta Y_{t+1} = 2\left(\alpha_2 + \beta_2 - 1\right),\ \Delta Y_{t+2} = 2\left[\left(\alpha_2 + \beta_2\right)^2 - \alpha_2 - 2\beta_2\right],\ ...$$

All responses should be scaled according to the size of the change in government expenditures. These algebraic expressions should correspond with the result from the computations made in the accompanying EXCEL book for specific examples. In stable models, responses of endogenous variables to a transitory change in an exogenous variable will go to zero relatively fast. Responses to a permanent shock in an exogenous variable will take endogenous variables gradually from their previous steady-state to the new one. In unstable models, in response to either a transitory or a permanent change in an exogenous variable, endogenous variables will permanently diverge. In larger scale models, characterizing the dynamics can be more complicated, since the *reduced form* equation explaining the behavior of an endogenous variable may well be of order greater than 2, as it was the case in the previous example. This is what happens in the model we discuss below.

It is important to bear in mind that multipliers are very easy to handle in linear models like the one we have considered. In models representing endogenous variables as implicit, nonlinear functions of exogenous variables, multipliers depend on the size of the change considered in the exogenous variables, and they may also depend upon the initial values from which the change is introduced. If the model is nonlinear, we cannot hope to solve anything similar to the characteristic equation, to give us the stability properties of the solution. The best we can do is to obtain the roots of the linearization of the model about a given point, preferable the steady state of the model, if it can be characterized. Unfortunately, stability of the linearized approximation does not guarantee stability of the original, nonlinear model. A second difficulty arises when actually trying to simulate the nonlinear model for given trajectories of the exogenous variables, as in the linear model above, since we will need to solve a nonlinear system of equations each period. As it is well known, even if it is *complete* such a system may have no solution, a single solution, or multiple solutions. Furthermore, the number of solutions may well depend on the range of values of the variables, so that what it is true one period regarding the nature of the solution, may not be true at some other points in time.

1.2.3 Stochastic, dynamic structural models

It is sometimes convenient to specify a stochastic model, in which we explicitly acknowledge that the behavior of each endogenous variable cannot be fully explained by that of the predetermined variables. In that case, we may include random perturbations as additional terms in some or all of the equations. These random variables will follow some specified probability distribution. For simplicity, it can be assumed that they are uncorrelated over time, as well as with each other, although this may not be fully realistic. That way, we would write,

$$\begin{aligned} C_t &= \alpha_1 + \alpha_2 Y_{t-1} + \varepsilon_{1t}, \\ I_t &= \beta_1 + \beta_2 \left(Y_{t-1} - Y_{t-2} \right) + \varepsilon_{2t}, \\ Y_t &= C_t + I_t + G_t, \end{aligned}$$

where ε_{1t} is the perturbation in the consumption equation, while ε_{2t} is the perturbation in the investment equation. We initially assume $E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$, $E(\varepsilon_{1t}\varepsilon_{1t-s}) = E(\varepsilon_{2t}\varepsilon_{2t-s}) = 0 \ \forall s \neq 0$, $E(\varepsilon_{1t}\varepsilon_{2t-s}) = 0 \ \forall s$, although we will later discuss how to cope with violations of some of these properties.

A shock to the consumption equation, i.e., a change in the value of exogenous innovation ε_{1t}, will have an impact on consumption this period, and also on output, through the aggregate income identity, with no effect on current investment. However, the increase in output at time t would have an effect on consumption, investment and output at time $t+1$ and beyond. An ε_{2t} shock will have an impact on current investment and output, but not on current consumption. However, dynamic effects will unfold from time $t+1$ on, as in the case of the ε_{1t} shock. These dynamic reactions are known as the *impulse response functions*, provided the shock takes place in a single period, i.e., that it is a purely transitory shock.

To actually compute numerically the impulse response functions, we start from the steady-state equilibrium values, with all the random perturbations in the model equal to the mean (zero), and assume that one of them takes for one period, a value equal to its standard deviation, with a positive or negative sign, depending on the type of shock we want to analyze. In addition to accumulating the impulse response function, if we want to compute the response to a permanent shock, we can also let the random perturbation to take a value equal to its standard deviation from time t on.

It is possible that a transitory shock may have contemporaneous effects on more than one endogenous variables. In the previous model, an ε_{1t} shock would have an impact on current consumption and also on output, through the national income identity. It will have an effect on future investment, but not on current investment. Similarly, an ε_{2t} shock would have a direct impact on current investment, and then also on current output and consumption.

That the random perturbations may present some autocorrelation is not hard to handle, since the equation can be quasi-differenced so that the transformed equation has an uncorrelated random error. For instance,

$$\begin{aligned} C_t &= \alpha_1 + \alpha_2 Y_t + \varepsilon_{1t}, \\ \varepsilon_{1t} &= \rho\varepsilon_{1t-1} + a_t, \end{aligned}$$

is equivalent to,

$$C_t = \alpha_1' + \alpha_2 Y_t - \alpha_2' Y_{t-1} + \rho C_{t-1} + a_t,$$

with $\alpha_1' = \alpha_1 (1-\rho)$, $\alpha_2' = \alpha_2\rho$, $E(a_t a_{t-s}) = 0 \; \forall s \neq 0$.

A more important difficulty arises when the random perturbations of the different equations are not uncorrelated with each other. We then need to introduce some identifying assumption. A popular method consists on establishing a rank of relevance among endogenous variables, using some ideas on causality. Then, if the random perturbation in the second equation in the ranking, is projected on the random perturbation from the first equation, the residual will be uncorrelated with the latter, and it can be interpreted as the part of ε_{2t} which is not explained by ε_{1t}. The random perturbation in the third equation could be projected on the random perturbations from the first two equations, and the residual would have a similar interpretation, and so on. To actually compute the impulse response functions, each equation in the model (except the first one), must be substituted by a linear combination of those that precede it in the ranking.[13]

1.2.4 Stochastic simulation

In previous sections we have seen how to simulate the model, generating time series of a pre-specified length for each of the endogenous variables. Necessary inputs for such a simulation are: values for the structural parameters, time series for each of the exogenous variables, as many initial conditions as lagged endogenous variables appear in the model and, in the case of a stochastic model, a time series for each of the exogenous random shocks. We will obtain a numerical value for each variable at each given period. However, we have not fully taken into account the fact that the random shocks in the model follow some specific probability distributions, or that we may have some uncertainty on the values of the parameters in the equations. These facts can be taken into account when performing Monte Carlo simulations.

For instance, to fully exploit the fact that the shock in each equation is a random variable, we simulate the model a large number of times, say

[13] Which is known as Cholesky identification strategy, from the way how a factor decomposition of the variance-covariance matrix of the original innovations is used to produce the linear transformation of the system of equations.

5000, sampling each time a different time series for each shock. The general approach to simulation consists on generating realizations for the stochastic shocks in the model, and use the model to produce stable time series realizations for all the relevant variables in the economy. That way, a probability distribution for the shocks in the model translates into a probability distribution for the vector of relevant variables. Given that distribution, characterized through a large number of simulations (numerical solutions), we will be ready to compute on our set of realizations, the values of any statistic of interest: *i*) output volatility , *ii*) relative volatility of consumption and investment to output, *iii*) correlations of consumption investment and interest rates with output, *iv*) cross correlations among any two variables, *v*) estimated coefficients in specific regressions, or *vi*) responses of a given variable to shocks in any other variable.

We will obtain a different numerical value for any of these statistics in each of the simulations we may run. If we ran 5000 simulations, say, we would obtain as many values of any of the mentioned statistics, so we will be able to approximate the probability distribution of that statistic through its empirical density. That way, we will be perfectly equipped to answer questions like: what is the probability that in this model, the consumption-output correlation takes a value below .92?

Uncertainty on parameter values can also be taken into account by specifying a priori a probability distribution gathering our beliefs on its possible values. For each simulation we would then use a different value for that parameter, chosen at random from its prior probability distribution. There are many probability distributions programmed in most statistical packages, so that almost any type of parameter uncertainty can be accommodated, to obtain simulations. We will just need to specify the numerical values of the parameters characterizing the chosen probability distribution. For instance, we could say that α_2 is Normal(0.85,4), and a numerical value sampled from this distribution can be used in each of the simulations. This is different from the case with no parameter uncertainty, in which the same value of α_2 would be used in all simulations. Parameter uncertainty makes sense when we *calibrate* a model (i.e., when we fix parameter values so that some implied statistics match their average values in time series data), or when the parameters are estimated by econometric methods. Theoretically, the number of simulations to be run should be increased to incorporate the fact that we should run for each parameter value, a large number of model simulations, all sharing the same numerical value for the parameter, but a different realization for the random perturbations.

It would be better to specify a single joint probability distribution for the parameters, as obtained, for example, from the estimation of a simultaneous equations econometric model. However, sampling from that distribution can be more complicated. Besides, if the model has not been previously estimated, the researcher may not have much information on the characteristics of that joint distribution. Nevertheless, the idea in Monte Carlo simulation is to specify as much information as we may have on the sources of uncertainty in the model in the form of probability distributions, to be used in simulation by drawing random realizations for each simulation from those probability distributions.

Even uncertainty over the paths of the exogenous variables can be taken into account this way: suppose we believe that, with probability p, government expenditures will increase at a rate of 1% every period over the simulation horizon, increasing at a rate of 2% with probability $1-p$. It would be sensible to run two different simulation exercises, with either path for government expenditures, to attach the mentioned probabilities to the resulting empirical frequency distribution for the endogenous variable being considered at a given point in time into the future. The researcher will then have two different empirical distributions for the value of that variable, each one having a given probability of occurring. Alternatively, a single Monte Carlo simulation exercise can be run, using one or the other path for government expenditures, with probabilities p and $1-p$. This way, we would have a single empirical distribution, possibly with two modes, reflecting the two alternative paths for government expenditures.

1.2.5 Numerical exercise: Simulating dynamic, structural macroeconomic models

EXCEL book *Dynamic responses.xls* shows simulation exercises for the dynamic models considered in the previous sections. The *Monotonic* spreadsheet considers a parameterization leading to a second order autoregression for output: $Y_t - .7Y_{t-1} + .1Y_{t-2} = .3 + G_t$, which is stationary, with roots .2 and .5. We consider an initial situation with government expenditures equal to 20 at all time periods, which leads to a steady state value of output of 50,75. We first analyze the effects of a one-period shock in government expenditures, that changes to a level of 21 at t^*, to return to the initial level of 20 afterwards. The output impact multiplier can be seen to be equal to 1, with negative dynamic multipliers afterwards that exactly compensate the initial response. The total long-run multiplier turns out to be zero. This must be the case in a stationary system, as we already know. The response may last longer than the initial shock, but it cannot be permanent. The second exercise looks at the effects of a permanent shock in government expenditures, which are assumed to jump to the level of 21 and stay there forever. The impact multiplier of output is again equal to 1, with positive dynamic multipliers, that make up for a total long-run response of 2.5. The graph to the left shows the output responses to a transitory as well as to a permanent shock in government expenditures. The graph to the right shows the output responses to a transitory shock in government expenditures in this and in the next model, which displays an oscillatory response, as we are about to see.

In the *Oscillatory* spreadsheet, numerical values for the structural parameters are chosen so that the second order autoregression for output is $Y_t - 1.4Y_{t-1} + .8Y_{t-2} = .3 + G_t$, whose characteristic equation has two complex conjugate roots $0.7 \pm 0.55678i$, with modulus of .8. That explains the oscillatory, damped cyclical responses that we see now to a shock in government expenditures. In the case of a permanent shock, the cyclical response takes the process to a new steady state for output, above the previous one, while the response to a transitory shock in government expenditures oscillates around the initial steady state for output.

The previous analysis has been performed in models without innovations. We have just changed the value of an exogenous variable, and examined the responses of endogenous variables to that shock. The *Stochastic G* spreadsheet considers a stochastic economy as in the last section, but with a single shock in government expenditures. In the spreadsheet we obtain a time series realization of 100 time observations for G_t out of independent $N(60, 3^2)$ random variable.[14] The equations of the model are used to obtain simulated data for the endogenous variables in the economy. First, we choose two initial values for output, Y_0, Y_{-1}, at its steady-state level.[15] The level of consumption at $t = 1$, C_1 is then obtained from the first equation, and the level of investment from the second equation. Since we already have the whole time sequence for government expenditures, we can now compute the level of output Y_1. Iterating on this scheme, we compute the whole time series for consumption, investment and output. To the right of output we have constructed time series for lagged output. Below the simulated time series data we see sample moments. Government expenditures have a mean of 60.19, with a standard deviation of 2.85. Average consumption is 92.22, with standard deviation of 2.39, average investment is 0.80 and average output is 153.20, with standard deviation of 3.98. Volatility is better indicated by the coefficients of variation, which is much higher for investment than for the other variables, a fact consistent with actual data.[16]

Consumption has a linear correlation coefficient of .69 with output, while the correlation of investment with output is lower, of .38. This model is so simple that it is easy to understand the nature of these relationships. From the first equation, the consumption time series has a unit correlation with lagged output, that has a correlation of .69 with current output. This is

[14] Alternatively, we could have considered a process with some inertia for Government expenditures, or even change the model to make the value of Government expenditures to be related to the past level of output, for instance.

[15] The choice of the steady-state level as initial condition is arbitrary. However, in this stochastic version of the model that choice is as good as any other choice, since the economy is already going to experience fluctuations due to the stochastic componente of government expenditures.

[16] Notice the difference between computing relative volatility by the ratios of standard deviations or through the ratios of the coefficients of variation, the latter option being preferible.

where the consumption-output correlation comes from. So, with the parameter values considered, the model introduces some persistence in output, as reflected on the correlation of .69 between Y_t and Y_{t-1}.This is also known as the first value of the autocorrelation function of output.[17] This persistence in output is possibly the more interesting feature for the model. It should be noticed that all these numerical values would change for a different realization of the stochastic process for government expenditures. They would also change if we change the stochastic process for government expenditures or any of the equations in the model, but also if we change the value of some structural parameter α_1, α_2, β_1, β_2. Changes in structural parameters will be important so long as they imply noticeable changes in the second order autoregression for output.

To continue illustrating the type of analysis that could be done out of simulated data, we may wonder about the type of consumption-output relationship emerging from this model. The model relates exactly *lagged* output to *current* consumption, but that is not the type of consumption function we are used to think about. The results of estimating such a consumption function, that relates current consumption to current output, are shown below the previous statistics. Because of the reasons already mentioned, we get some explanatory power, with a R^2 coefficient of .48, and an estimated slope of $\hat{\beta} = .42$. The first graph below displays residuals as a function of the explanatory variable, output, with no much evidence of relationship. The graph below shows them as a function of the dependent variable, showing a positive relationship, consequence of the fact that there is a significant component of consumption that remains unexplained by the regression on output and it is therefore included in the regression residuals. The first graph to the right shows residuals as a function of time, with no evidence of persistence. Residuals can be seen to cross their mean value of zero very often. Finally, the graph below shows the consumption-output scatter diagram and the fitted regression line. Time series for the fitted consumption values and the implied residuals are shown to the right of the time series for endogenous variables. Lagged residuals are also displayed and the first order autocorrelation coefficient of .11 is presented at the end of the series.[18] We have included a second spreadsheet *Stochastic G* (2) differing from the previous one only in the sample realization for government expenditures, so that the reader can see what changes can be seen in numerical values of the different statistics as a consequence of the stochastic nature of the model.

The *Multiple shocks* spreadsheet repeats the exercise, this time considering innovations in the consumption and the investment equations, as well

[17] The autocorrelation function is the sequence of values $Corr(Y_t, Y_{t-s})$, for all s.

[18] This suggests no evidence of residual autocorrelation, a potencial source of misspecification in the consumption equation.

as in the stochastic process for government expenditures. The process for government expenditures is the same as in the previous spreadsheets. The linear correlation coefficients of consumption and investment with output are now higher than in previous exercises. This is due to the fact that the consumption innovation affects both, the level of consumption and also the level of output at each time period, so that there is a common stochastic component. The same argument explains the higher correlation between investment and output.

The *Impulse responses* spreadsheet computes responses to transitory and permanent shocks in each of the endogenous variables: consumption, investment and output. These responses are obtained as follows: initially, all variables are supposed to be at their steady-state levels. All innovations take a zero value, so that at all effects it is as if we consider a deterministic model. At some time $t = 0$, an endogenous variable takes a value equal to its steady-state level plus an increase (the impulse), of size equal to one standard deviation, and we compute how all variables evolve from then on. For the size of the impulses, we take standard deviations form the stochastic version in the version of the model when only government expenditures were random.[19] Consumption and output are shown to react strongly to an impulse in consumption. Investment reacts with a one period delay, and the response is very short. Impulses on investment do not have much effect on either consumption or output. Consumption and investment show a strong response to output shocks with a one period delay, the response of investment extending to just one period.

The two previous sections have allowed us to introduce statistical concepts that will be used throughout the book when analyzing numerical solutions to Growth models. We have also advanced some of the fundamentals of Monte Carlo simulations of dynamic models, to show how the statistical and econometric analysis of the set of time series obtained as solution to the model allows us to deduce a much richer set of implications than could be obtained analytically. We now move into describing the main characteristics of Growth models, their evolution following a variety of research interests, how they are equipped to deal with Lucas' criticism on policy evaluation, and how their numerical solutions can be obtained and exploited for policy analysis.

---

[19] This is arbitrary. We should take an impulse of size equal to one standard deviation of the innovations estimated from actual time series data, since that is the likely single-period fluctuation in each variable.