The business cycle: dynamical coupling and chaotic fluctuations

de Langen, F.

Citation for published version (APA):
CHAPTER 2

THE BUSINESS CYCLE: EXPLANATIONS AND CONTROVERSIES

"Any answer in terms of a single cause is sure to be wrong" (Schumpeter(1982), 34)

1 Introduction

There are significant theoretical controversies regarding both the dominant causes and the economic mechanism which generate fluctuations. This chapter highlights the controversy between the endogenous and exogenous business cycle theories. The endogenous approach is defined here as those theories that assume the economic system to exhibit persistent fluctuations caused by endogenous relationships. Exogenous cycles are based on the assumption of a stable, but disturbed, economy, which only displays fluctuations in the presence of exogenous shocks.

The impulse-propagation problem
Frisch(1933, 171) remarks: "we have to distinguish between two fundamental problems: first the propagation problem; second, the impulse problem". The propagation problem describes the structural dynamical characteristics of the system, the impulse problem treats the initial disturbance which distorts the system. An impulse can only lead to persistent fluctuations if the underlying economic structure supports those fluctuations (also see Haberler(1939) and Kyun(1988, 30) who describes it as Pigou's propagation mechanism). Frisch(1933) concludes that the ‘dynamic laws of economics’ must show dampened oscillations, founding this theory on the examination of different dynamic economic models and the state of mathematics. It is, therefore, necessary to assume “a stream of erratic shocks that constantly upsets the continuous evolution, and by doing so introduces into the system the energy necessary to maintain the swings” (Frisch(1933), 197/198). His explanation of the impulse is based on Wicksell and Akerman (Det Ekonomiska Livets Rytmik, Lund, 1928, see Frisch(1933)). The latter used an analogy of the business cycle as a stream of water flowing in an uneven riverbed with irregular waves. Although Frisch criticised this approach, the theory on waves is, at present, one of the subjects of the research on persistent (chaotic) fluctuations. Some of this research has found applications in economics.
**Exogenous versus endogenous theories**

The division between the original impulse and the propagation mechanism can be used to clarify the classification of business cycle theories. One set of theories of business fluctuations relies on the regular\(^1\) occurrence of shocks, the impulses. These theories are labelled *exogenous theories*. Within this category two approaches are possible. One branch of the exogenous approach assumes the economy to adjust immediately to shocks. To explain fluctuations, it has to assume impulses to behave in a wave-like manner. A second branch assumes, as Frisch, a temporally adjustment process, which could take the form of fluctuations. These fluctuations vanish in time, so the continuation of fluctuations also requires a continuation of disturbances. In this approach, the explanation of the impulses is as important as the explanation of the propagation mechanism.

Other explanations of the business cycle concentrate on the structure of the economy (the propagation mechanism) to explain a lasting cycle: the *endogenous theories*. To do so, certain lags between the key-variables, independent of the original impulse, are assumed to exist. Haberler(1939, 11) states: “*historical evidence seems to demonstrate that the cyclical movement has a strong tendency to persist, even where there are no outstanding extraneous influences at work which can plausible be held responsible*. .. Exogenous forces will then figure as the originators of disturbers of endogenous processes, with power to accelerate, retard, interrupt or reverse the endogenous movement of the economic system". These theories give primacy to the propagation mechanism, without denying the occurrence of external disturbances that initially cause the economy to deviate from its equilibrium.

This view differs from that of Gandalfo(1996, 531): “*the distinction between endogenous aperiodic behaviour and exogenous aperiodic behaviour is slippery in applied work, since -generally speaking- it depends on the size of the model. When the small model is considered as part of a larger model, some of the exogenous influences [stochastic exogenous disturbances] may be taken in as endogenous variables*. Theoretically, there is a difference of opinion on the structural characteristics of the economy (Frisch versus Haberler), although this difference is not always seen as important in applied work (Gandalfo).

The distinction between endogenous and exogenous fluctuations is relevant for two reasons. First, this distinction is reflected in the empirical work on the business cycle. If the presumption is that fluctuations are caused by exogenous disturbances, the aim of the analysis will be to identify those disturbances and their relationship with economic variables (Gandalfo’s small models). If, however, an initial disturbance causes ongoing

---

\(^1\)To generate irregular time series in a stable model, the assumed shocks have to be regular in appearance, but irregular in magnitude.
fluctuations, there is reason to search for economic explanations of the successions of booms and depressions (Gandalfo’s large models). The conclusions from empirical work have implications for the second line of research, that on the influence of the government on business fluctuations. If the economy is inherently structurally stable, the main purpose of government policy is to maintain a stable environment with as little distortions as possible. It is only when periods of adjustment are too long, that there is a reason for interventions.

When the economic relations lead to persistent fluctuations, there is a role for intervention. By changing its own behaviour, the government can influence the outcome of the economic process into a desired direction. In this case it is important to have knowledge of the source of potential disturbances and the nature of the propagation mechanism.

Real versus monetary
Another source of discussion is the nature of the impulse and propagation mechanism. Milton Friedman, in correspondence with Snowdon, Vane and Wynarczyk, declares: "Disturbances of all kinds may produce fluctuations in economic activity. Those disturbances may be monetary or they may be real. There is no reason why they should be consistently one or the other. As a matter of experience, monetary disturbances seem to predominate, though other disturbances have been important from time to time" (Snowdon, Vane and Wynarczyk(1994), 175). Some, as Friedman, stress the monetary structure as the main source of instability and fluctuations, others give priority to the real part of the economy. A third group of economists seeks an explanation for the fluctuations in a combination of both.

As the quotations of Friedman and Schumpeter show, almost all economists accept the presence of several kinds of disturbances. There have been major disputes regarding the dominance of real shocks over monetary disturbances and vice versa. This, however, is secondary to the discussion on the inherent stability of the economy: does the propagation mechanism show persistent fluctuations after a one-time-disturbance or will the economy return towards some kind of equilibrium?

2 Exogenous explanations of the business cycle

Most of the economic thinkers in the eighteenth century did not present an explicit theory of the business cycle. Only after the major economic fluctuations in England and France, after the Napoleonic Wars (1793-1813), a discussion on economic dynamics emerged. Depressions and booms were assumed to be temporary and independent phenomena, not the result of regular fluctuations. The occurrence of regular seasonal fluctuations was easily explained since, as Blaug(1998, 15) observes, “in seventeenth-century England, a predominantly rural economy, most unemployment was due to the seasonal nature of agriculture or to the incidence of poor harvests. Even
in industry, much unemployment was seasonal, as winter ice or spring
floods interrupted the functioning of the water-powered mills". Classical
economists such as J. Mill and J. B. Say explained the occurrence of the
depression as part of the adjustment process after an exogenous shock: the
economy moves from one equilibrium position to another.

Later in the nineteenth century other economists became interested in the
relationship between periods of recovery and depressions. One of them was
W. S. Jevons (see Duindam(1992)). His business cycle theory started from the
statistical observation of the similarity between economic fluctuations and
the periodic occurrence of sunspots. Sunspots influence the average
temperature, which determines the size and quality of harvests.
Fluctuations in the supply of agricultural products did cause major
fluctuations in the costs of living (and the wages), production of industrial
products and trade. More sophisticated measurement of the sunspot cycles
and the declined importance of the agricultural sector have undermined the
relevance of this theory.

The term 'sunspots' is still in use in economics. Cass and Shell (1983) and
Azariadis and Guesnerie (1986) have done initial work on sunspots.
Azariadis(1993, 426) uses an economy described by a general function to
illustrate the occurrence of sunspots:

\[ x_{t+1} = f(x_t, \mu_t) + \epsilon_{t+1} \]

With:

- \( x_t, x_{t+1} \) = the state vector at \( t \) and \( t+1 \)
- \( \mu_t \) = parameter vector, representing preferences, production
  possibilities, initial endowments and population, which can
  change in time.
- \( \epsilon_t \) = an arbitrary random variable, i.i.d. on \([-a, a]\) around zero, with
  zero mean

Azariadis(1993) then assumes a stationary deterministic structure with \( \mu \)
constant (given preferences, production possibilities, population and initial
endowments). "Any randomness remaining in the solution [of the equation
above] must come from extraneous or psychological factors for which David
Cass and Karl Shell(1980) have invented the convenient name "sunspots".
Further on, Azariadis(1993, 477) remarks: "There are all sorts of
intermediate possibilities whose realization depends on public beliefs, that
is, on what variables individuals consider to be salient for economic
prediction... many forecasting rules turn out to be consistent with rational
expectations and each one supports an equilibrium that allocates a distinct
random consumption bundle to each consumer. We call these allocations
payoff-relevant sunspot equilibria... sunspot allocations are Pareto
dominated by the corresponding non-sunspot outcomes because the
sunspots inject spurious noise in an economic environment whose fundamental structure may very well be quite stable".
The role of expectations is discussed later in this chapter. In terms of Frisch, $f(x, \mu)$ is the propagation mechanism and $\varepsilon$ the impulse.

The importance of exogenous shocks carried over into the post-war economic theories. The sources of fluctuations were not at the centre of the disagreement, rather it was the possibility of the government to intervene, to stabilise the economy. The dominant economic theories in this period were the Keynesian and the monetarist approaches. The monetarist stance in this debate was mainly the adherence to the long-term stability of the economy. Government policy is at its best ineffective, at its worst pro cyclical.

**Early Keynesian business cycles**

Building on the work of J. M. Keynes in "The General Theory of employment, interest and money" (1936), Samuelson (1939), J. Hicks (1950) and R. Goodwin (1951) developed different versions of the multiplier-accelerator models, using a relationship between income and fixed capital. Other variations of the accelerator-multiplier model, as for example, Metzler (1941) proved that inventories could play the same role as investment in fixed capital. Inventories are held to smooth out production, according to differences in sales over time (see Gabisch and Lorenz (1989), 53). Fluctuations in these models depend on the structure of the economy, as modelled by the lagged reaction of consumption and investment to changes in disposable income. The different multiplier-accelerator models take a variation of the following formulation:

$$Y_{t+1} = \alpha Y_t + \beta Y_{t-1} + \chi Y_{t-2} + A_0$$

With:

- $Y$ = (national) income
- $t$ = period
- $A_0$ = autonomous spending
- $\alpha, \beta, \chi$ = parameters

The 'multiplier-accelerator'-economy exhibit several kinds of dynamics, depending on the parameters ($\alpha, \beta$ and $\chi$) as given in table 2.1.

<table>
<thead>
<tr>
<th>Cyclical</th>
<th>Monotone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dampened</td>
<td>Cyclical towards $Y^*$</td>
</tr>
<tr>
<td>Stable</td>
<td>Regular cycles around $Y^*$</td>
</tr>
<tr>
<td>Explosive</td>
<td>Explosive cycles</td>
</tr>
</tbody>
</table>

Table 2.1: The dynamics in the multiplier-accelerator model ($Y^*$= equilibrium position).

Explosive behaviour is ruled out as it is not observed in reality. The occurrence of regular cycles depends strongly on specific parameter constellations. Therefore, it was initially assumed that a 'multiplier-
accelerator'-economy would exhibit dampened dynamics. In this case the occurrence of business cycles depends on exogenous shocks in $A_0$ and is labelled exogenous here. When the model is adjusted, endogenous cycles are possible. This case will be treated in the next section.

**New Classical business cycles**

When growth declined in the seventies and especially after the failure of Keynesian demand policy to provide a remedy for the depression following the first oil price shock in 1974, both governments and theoretical economists returned to the classical theories of the business cycle. Theoretically, the Keynesian macroeconomics was criticised for the absence of a microeconomic foundation. A new approach was developed building upon classical microeconomic insights. This set of models became known as the new classical economics, because of the emphasis on the market clearing feature of prices and the optimising behaviour of consumers and producers. It is also called the rational expectations revolution due to the assumption of the rational formation of expectations, in the sense of Muth\(^2\). A problem for the new classical economist was the explanation of fluctuations (see Lucas(1985), 7). When economic subjects know both the data and the true structure of the economy, there is no reason for the economy to deviate from its long-term growth path. Two different explanations were developed from the equilibrium approach.

**Rational Expectations**

R. E. Lucas introduced imperfect information in a monetary rational expectations model. As was concluded in his 1979-article\(^3\): "This paper develops a theoretical example of a business cycle, that is a model economy in which real output undergoes serially correlated movements about a trend which are not explainable by movements in the availability of factors of production. The mechanism generating these movements involves unsystematic monetary-fiscal shocks, the effects of which are distributed through time due to informational lags and an accelerator effect. Associated with these output movements are (i) pro cyclical movements in prices, (ii) pro cyclical movements in the share of output devoted to investment, and (iii), in a somewhat limited sense, pro cyclical movements of the nominal interest rates. This behaviour is obtained under the assumptions about expectations formation which seem suited to the study of a recurrent event: agents are well aware that the economy goes through recurrent 'cycles' which distort perceived rates of return. On the other hand, the transitory nature of real investment opportunities forces them to balance the risk of incorrectly responding to spurious price signals against the risk of failing to respond to meaningful signals."

\(^2\)Zijp(1992) gives an analysis of the differences between the new classical economics and the Austrian economics, also treating the formation of expectations.

Lucas uses the idea of Phelps-islands (see also Chari(1999)). Each producer lives on an island (its own industry or branch). The producer has good knowledge of the situation of its own market, but imperfect information on the total economy. Assume the government to finance its spending by the creation of new money. This stimulates nominal demand and raises the price level, on a macroeconomic level. The producer observes only a rise in demand for its own good. The decision to be taken is that of whether the observed rise in demand is a rise in (real) market specific demand or a (nominal) rise in the general price level. Depending on the historical fluctuations in the general price level, the producer chooses the amount of production. In a non-inflationary environment, the rise in nominal demand leads to an expansion in production. The rise in production depresses prices, since real demand stayed constant. As more information becomes available in time, it becomes clear that the rise in demand was a general nominal rise. The decline in nominal demand, however, again surprises some producers, causing production to decline. Depending on the structure of the economy, the initial rise in the money supply generates dampened fluctuations, when expectations are adjusted to experience. Of course in a historically inflationary economy, a rise in nominal demand has little or no impact. The learning-effect causes the equilibrium approach to the business cycle to conclude that systematic monetary policy is ineffective.

Real business cycles
The new classical approach depends strongly on the reoccurrence of monetary shocks and imperfect information. Given a learning process of the maximising individuals and the faster availability of information on monetary aggregates, some economists viewed these assumptions as unlikely. Keeping the equilibrium concept central, attention shifted towards real shocks, causing the real business cycle approach to emerge. The real business cycle approach seeks the impulses of the business cycle primarily in real changes in technology, as measured by the Solow-residual, or preferences of the consumers; no distinction is made between growth and cyclical fluctuations. This approach emphasises the intertemporal substitution between leisure and consumption on one side and the natural rate hypothesis on the other. The fluctuations are the result of adjustments, as the economy moves from one equilibrium growth path to another and the accompanying smoothing of consumption over time (as was the belief of the classical economists). Adjustments are time-consuming because the time to build-mechanism. Kydland and Prescott(1982) give an early example of this model (see Blanchard and Fisher(1989), chapter 7 for other examples). Here, we follow King, Plosser and Rebelo(1988, I and II), from here KPR(1988I or 1988II), in their description of the basic real business cycle model. They assume many infinitely lived individuals to maximise a concave utility function with leisure and consumption as goods. In order to ensure a feasible steady state, the following restrictions are made:
A. on production: "permanent technical change must be expressible in a labor augmenting form" (KPR(1988I), 200).

B. on preferences:

i- the intertemporal elasticity of substitution in consumption is invariant to the scale of consumption. The marginal product of capital equals one plus the real interest rate, and is constant in the steady state. The marginal discounted utilities also equal one plus the interest rate, so "the intertemporal elasticity of substitution must be constant and independent of the level of consumption" (KPR(1988I), 201);

ii- the long-term income and substitution effects of sustained growth in labour productivity do not change the labour supply, because hours worked in the steady state cannot grow.

Combined, these assumptions imply that -given the initial ratios to be the result of utility- and profit maximisation- the levels in a new equilibrium change but their comparative ratios stay constant.

Production takes place using capital and labour. Markets are assumed to clear because of the adjustments in (flexible) prices. Temporary shocks in technology are modelled as changes in total factor productivity, whereas permanent technological changes will take place through changes in labour productivity. KPR(1988I) shows the existence of a general growth equilibrium, assuming labour productivity to grow at a constant rate. The path of efficient capital accumulation is approximated by:

\[
\dot{k}_{t+1} = \mu_1 \dot{k}_t + \psi_1 \dot{A}_t + \psi_2 \sum_{j=0}^{\infty} \mu_2^{-j} \dot{A}_t + j + 1
\]

With:

- \(K_t\) = capital stock
- \(X_t\) = level of labour productivity in period \(t\)
- \(\dot{k}_t = K_t / X_t\)
- \(\dot{k} = \log(k_t / k)\)
- \(k = \) equilibrium value of capital, in terms of \(X\)
- \(\dot{A}_t = \log(A_t / A)\) = percentage deviation of actual total factor productivity (\(A_t\)) from equilibrium total factor productivity (\(A\))
- \(\psi_1, \psi_2 = \) functions of the preferences and technology
- \(\mu_1, \mu_2 = \) coefficients

As KPR(1988I, 206) state: "The dynamics of capital accumulation depend on the previous period's capital stock with coefficient \(\mu_1\). In addition, with time-varying total factor productivity, the optimal solution for capital...

\[4\]KPR(1988I, 206), equation (3.7), symbols as in KPR(1988I).
accumulation depends on the current productivity level \( (\bar{A}_t) \) and on the entire future time path of displacements to productivity ‘discounted’ by \( \mu_2^". \)

To satisfy the transversality condition it is necessary for \( \mu_1 \) to be smaller than \( 1^5 \). Equation (2.1.1) then shows a monotone movement towards the steady state. A persistent shock to technology does only influence the level of the equilibrium growth, not the relative proportions.

Three propagation mechanisms explain the occurrence of cycles:
1. ‘Time to build’: the intertemporal effect of capital accumulation.
2. The intertemporal substitution of leisure: the trade-off between leisure and consumption. This mechanism explains the difference in reaction of labour to a permanent and a temporary technological shock. A rise in income has a substitution-effect: the opportunity costs of leisure rise, so there is an incentive to provide more labour. Also, there is an income-effect: income rises as the same amount of labour is provided, inducing the supply of labour to fall. The total effect of a rise in labour productivity (and the accompanying rise in wages) depends on the relative weight of both effects. When the shock is perceived as permanent, the substitution effect is nil so on average the short run labour supply declines. If the rise in income is seen as temporary, the substitution effect is high and labour supply rises.
3. the interest rate or wealth-effect: as total productivity rises, so does the interest rate. An increase in the interest rate decreases the discounted value of future wages (wealth) and makes it more attractive to work today. Individuals accumulate wealth (capital), inducing them to work additional hours, despite the low real wage rate. When wealth rises towards its equilibrium value (in the absence of new shocks) consumption will catch up, real wages rise and hours worked decline.

Comparing the conclusions of the model and data on the United States KPR(1988I, 231) conclude that “the basic neoclassical model is capable of replicating some stylized facts of economic fluctuations”, so “the principle serial correlation in output -one notable feature of economic fluctuations- derives mainly from the persistence of technology shocks” (In a sense this conclusion is foreseen by Frisch(1933), 198/203).

In KPR(1988II) extensions of the basic model are analysed (stochastic technological change, endogenous technological change, sub optimal equilibria and heterogeneity in consumers). The conclusions of the basic model do by-and-large carry over to the more complicated models. The debate between the real business cycle theorists and other economists is still active. Snowdon, Vane and Wynarczyk(1994) give an overview of this discussion. One conclusion of the real business cycle approach makes this approach very vulnerable to criticism: because of the reliance on the

\[ 0<\mu_1<1<\beta^{-1}<\mu_2, \] with \( (\mu_1, \mu_2) \) the roots of the first-order dynamic system in \( k \) and \( \lambda \) (the shadow price of capital) and \( \beta \) the rate of time preference, see KPR(1988I, 203/204).
intertemporal substitution between consumption and leisure, all unemployment is due to the consumer’s free choice. This might be true for some individuals, especially in prosperous times, but it is doubtful that depressions are caused by the preference for leisure of large groups in society.

3 Endogenous approaches to the business cycle

In the same period as the classical economists, other economists such as J. Sismondi de Sismondi and T. R. Malthus developed the under-consumption theory, which sought to give an endogenous explanation of depressions. The under-consumption theory is not a true business cycle, in the sense of an explanation of both depression and booms. As with the exogenous classical theories, growth was seen as the ‘natural’ state of the economy. Neither author, however, depends on exogenous shocks for the economy to leave its equilibrium growth path, but endogenous influences cause the economy to go into a depression.

Malthus and Sismondi see the unequal distribution of wealth as the primary cause of the crisis. Malthus assumes the owners of capital and land to save more than the wage earners. In a growing economy, the wealth of land- and capital owners grows faster than wage income, increasing savings and depressing consumption. The optimal ratio between savings and consumption is disturbed, and overproduction results. Demand for consumption goods declines but the production capacity rises with the savings. This results in a depression, with declining prices and profits. The decline in profits enlarges the depression. The French economist Sismondi emphasises the complexity and lack of planning in the capitalist economy. Due to this, producers have only one source of information: present prices. Current prices determine future supply, since the behaviour of the competitors and the potential buyers is unknown. In this sense Sismondi can be called a forbearer of the new classical economists. Secondly, current income, which determines current consumption, is production from the former period. When the economy is growing, present consumption is always smaller than present production: there is an ongoing tendency toward overproduction (under-consumption). This results in periodic crisis. Prices only adjust slowly because of the imperfect information. Only after a depression, in which the distribution of wealth adjusts towards its old proportions, growth returns. As before, growth causes overproduction, a crisis and a depression. Again, growth is the normal state of the economy, only negative deviations of this trend have to be explained.

In the same period J. Mills (Credit cycles and the origin of commercial panics, 1867, see Mullineux(1990), 64) published his theory of the psychological cycles. According to Mills the business cycle is monetary in nature. The economic cycle is caused by the cycle in the psychology of
businessmen. Depending on the psychological state of the entrepreneur and the banks, entrepreneurs ask for credit and banks will supply or deny it. The psychological state behaves in a cyclical way. Starting with a panic, this leads to the refusal of credits and the withdrawal of loans. The following period of depression sees a restructuring of the economy. Less profitable investment is written off and firms without sufficient profit opportunities go bankrupt. This raises the average profitability, confidence builds up and credit opportunities become larger. This optimism creates a boom. Eventually, the optimism will lead to much risk taking, causing business failures, a downfall and a panic. Although the psychological state of the businessman is outside the domain of economic reasoning, the interdependency between behaviour and economic state makes us categorise this theory with the endogenous business cycles. This economical-psychological explanation of the business cycle became part of many modern theories.

3.1 Disproportionality and the business cycle

The interdependency of booms and recessions became recognised by many economists, as can be seen from the work of Jevons and Mills. The business cycle was seen as the result of a deviation between the structure of production and the structure of demand (the disproportionate production argument) as described in the theories of Malthus and Sismondi. Business cycle theories try to give an answer to the cause of this disproportionality.

Haberler (1939) distinguished three kinds of economic decisions:
1. the decision between saving and spending of income;
2. decisions about the distribution of spending regarding different consumer goods;
3. decisions of producers about the distribution of inputs regarding the different stages of production.

When the structure of production does not match the preferences according to consumption and saving (the first economic decision), Haberler speaks of vertical maladjustment. This is caused by the difference between saving and investment.

It is also possible that the structure of production deviates from the desired structure of consumption (the distribution of demand over different kinds of consumption goods, because of decisions of the second and third kind). This is horizontal maladjustment: the producers invest in the production process (investment equals savings), which produces goods that are not in demand. Other goods are in excess demand, due to the misdirection of funds in production.

Vertical maladjustment

M. Tugan-Buranovsky and A. Spiethoff are two early representatives of the first approach. The fundamental source of fluctuations lies in the difference between saving and investment, causing changes in the structure of
production of capital goods. Although monetary factors play a part in their business cycle theory, it is only a passive secondary role. The upswing starts when investment rises, as funds are accumulated during the depression and new profitable investment opportunities arise. The spending on capital goods and ‘investment goods’ (inputs for the production of capital goods) increase income. This rise in income stimulates further increases in investment, but also raises the prices of capital- and investment goods. Profits rise causing a greater availability of (internal and external) funds and an emerging optimism. Seemingly there is no reason why this expansion will not go on, but “it is an actual shortage of capital that brings about the crisis .. capital shortage does not mean simply a deficiency of monetary funds, but is the symptom of a serious disproportion in the production of certain well-defined types of goods .. over-consumption which leads to a scarcity of capital and brings about the end of the boom” (Haberler(1939), 75).

During the upswing wage-income rises as the demand for labour rises. The propensity to save out of wages is below the savings rate of profit income. The demand for producers’ goods declines and the demand for consumers’ goods increases at the end of the upswing: “they [the capital goods industries] will be depressed although, or even because, the consumers’ goods industries prosper” (Haberler(1939), 78). The sector producing consumers’ goods can only slowly expand as production factors (money, labour and materials) are used by the increased production of producers’ goods. Savings are too low (over-consumption) to finance the expansion in the production of producers’ goods. Both producers of capital and consumption goods will compete for additional credit, which raises the market interest rate. Profitability declines in the producers’ goods sector, which will depress investment further. The actual downswing is characterised by “[P]essimism and reluctance to invest and to embark on new enterprises .. The process of contraction also has a cumulative nature. Pessimism and reluctance to invest cause a shrinkage in the volume of purchasing power .. everything that increases the rate of savings (e.g., inequality in the distribution of income) has a depressing influence .. Prices fall and this intensifies the prevailing pessimism” (Haberler(1939), 79-80). Pessimism, hoarding of money, restricted credit and rising interest rates cause a restructuring during the downswing. The severity of the downswing depends among others on the way the firms and banks react to the initial crisis. When there is a large financial reaction and many bankruptcies the downswing will be large. When the downward adjustments take place at a slower pace, only a relatively small depression will result.

“The state of depression is interrupted (a) because it creates automatically a situation favourable to the revival of investment [large accumulation of funds, declining interest rates and wage rate], (b) because pessimism disappears with the lapse of time, and (c) because of the introduction of
stimuli from outside [new markets, new inventions (in the sense of Schumpeter)]" (Haberler(1939), 84). This causes a new upswing. An additional argument for the occurrence of the upswing is the need for replacement. Assuming the majority of capital goods to be installed during the same period, some of it has to be replaced in a future period, starting a new investment boom.

**Horizontal maladjustment**

Theories of the second approach, the horizontal maladjustment approach, assume equality between savings and investment. The cyclical fluctuations are caused by a lag between investment and the actual production. Adjustment of the capital stock to the expected final demand is time consuming. Production (emerging at the end of the production period) is based on expectations and observations at the beginning of the production period. This can be illustrated using a simple model based on the models of M. Kalecki (see *A Macro-dynamic Theory of Business Cycles* (1935), reprinted in Kalecki(1971)).

![Diagram of Kaleckian business cycle](image)

Figure 2.1: The Kaleckian business cycle; A gives the cycle of the production of capital goods, D the installation of investment and I the orders of capital goods. The time between ordering and instalment is \( \theta \).

Kalecki(1971, 10)

The Kaleckian economy is divided into two sectors producing capital goods and consumer goods. In society there are two classes, capitalists, who save their total income (profits) and workers, who consume their total income. When the capitalists who produce the consumption goods expect a rise in future demand, they invest. They order more capital goods than needed for replacement (depreciation). Production and instalment of the capital good take one period. So the orders in period \( t \) become productive in period \( t+1 \).

---

6Kalecki(1971) estimates the following cycles:

- \( I-U = a \sin 0.63t \)
- \( A-U = 0.98a \sin 0.63(t-0.3) \)
- \( L-U = a \sin 0.63(t-0.6) \)

With \( U = \) (constant) depreciation, \( L = \) deliveries, \( I = \) investment orders and \( A = \) output of capital goods. Note the terminology used by Kalecki is different from the one used above.

The Business Cycle
During the production process, labour income and profitability rise. The rise in profitability causes net investment in the second sector to rise. The rise in labour income stimulates demand in the consumption good sector. Prices of consumption goods rise because supply is given, as the ordered capital goods are not yet installed. This increases profitability in the consumption goods sector, again advancing orders for capital goods (investment).

In the next period, the capital goods are installed in the consumption goods sector. When the production increases above demand, the price of the consumption good declines, so the desire to invest also declines. Delivery of capital goods depends, however, on orders in the former periods. The addition to the capital stock goes on for some periods. Due to the decline in orders for capital goods income in this sector decreases. Demand for consumption goods declines and lowers investment. Orders for capital goods from the consumption goods sector decline below depreciation. Production capacity also declines. In time, this causes the supply of consumption goods to decline below demand, prices are rising and again orders are placed for new capital goods: starting a new cycle. The short-term fluctuations are caused by the lag in adjustment between the supply of and demand for consumption goods, which is based on the lag between the ordering and installation of the new capital good.

The business cycle theory of Kalecki shows similarities with the multiplier-accelerator models of Samuelson(1939) and Hicks(1950) (see paragraph two): investment is the leading force in the economic dynamics. Whereas Kalecki viewed investment as determined by the level of profits, the Keynesian models saw changes in the capital stock as a reaction to changes in demand. In the former section we labelled the linear multiplier-accelerator models as exogenous, because regular cycles are only possible for specific values of the parameters and the observed fluctuations are less regular in amplitude and period, so only dampened fluctuations are likely to occur. Hicks(1950) combined growth and the possibility of a larger variation in the parameters. The propagation mechanism is as described previously. The accelerator is assumed to take a value that gives explosive fluctuations. The fluctuations do not destroy the economy because production is restricted from above by the maximum production possibilities of the economy (given the available resources). The lower limit of change is given by autonomous investment minus depreciation. When investment reaches the floor or ceiling, only the multiplier-effect is active, stabilising the system. These models are known as floor-ceiling models (see Gabisch and Lorenz(1989), 60-62). Goodwin(1951) also describes a non-linear accelerator model.

Later work on non-linear multiplier-accelerator models showed the possibility of chaotic behaviour. Extensions of the Goodwin and the Kaldor model are given in Gabisch and Lorenz(1989). For a chaotic version of the Hicks-model, see Hommes(1991, 1993). Both the floor-ceiling models and the chaotic versions of the multiplier-accelerator model can be labelled endogenous theories.
3.2 Other real endogenous business cycles

Predator-prey models
Another approach originates with Richard Goodwin (Goodwin(1967), also see Goodwin(1987) for an historical review). The predator-prey model is based on biological models modelling the war of existence between two populations. The two competing species in Goodwin’s model are labourers and capitalists. The labourers receive a wage income. The remainder of production is profit income for the capitalists. In the simple version of the model, total wage income is consumed and profit income is saved. The Phillips curve determines the wages: a high level of unemployment causes the wages to decline. Unemployment results as the growth in demand for labour (determined by the rate of capital accumulation and growth in labour productivity) is below the growth in labour supply (determined by growth in population). Growth in labour productivity and population are assumed exogenous, capital accumulation is determined within the model. When profits rise, capitalists invest, so the capital stock grows. As the growth of capital (and labour demand) exceeds the growth of population, wages rise and profits decline. The period of expansion (fast growth in output) comes to a stop. Capital accumulation decreases, depressing the demand for labour. A period of depression begins (lower growth of output). The declining capital accumulation leads to unemployment. This depresses the wage share in income, restoring the profits. Slowly, investment will pick up. The accompanying rise in production causes a new expansion to start again. During the expansion employment and the wage share rises. Wages prey on profits, as the lynx on the hare in Hudson Bay (Rosser(1991), 237).

Investment depends on the profit income, as in the Kalecki-model, whereas investment in the multiplier-accelerator model depends on demand. Cycles in predator-prey models exist for more parametric values than in the multiplier-accelerator models. Given the initial position, each parameter configuration in which ‘the predator does not kill all the prey’ is stable (see Rosser(1991) on the equivalence between economics and ecological systems).

Modern Keynesian models
The critique on the micro-foundation of the traditional Keynesian models led to a reinterpretation of Keynes’ theory. Two approaches were developed: the neo-Keynesian theory (see Malinvaud(1977)) and the new-Keynesian approach (see Mankiw and Romer(1992), vol. 1 and 2). Business cycle theory is never been at the core of those two approaches. Their main goal is to give a micro-economic explanation of the short-term inflexibility of prices and wages.

These models were introduced independently by Volterra and Lotka, hence the Lotka-Volterra model. See Gandalfo(1980, 448-481), for a mathematical treatment.
Benassy (1986) developed a business cycle theory, based on the neo-Keynesian model. He assumes a Phillips curve, so prices and wages adjust slowly when notional demand and supply are not balanced. Trade equals the minimum of effective demand and effective supply. Investment is made endogenous assuming a structure similar to that in the multiplier-accelerator model. The starting point is a Walrasian equilibrium where notional demand and supply on the goods and labour market coincide. Presume a positive demand shock, so demand for goods increases. Prices and wages remain constant in the short run. Firms would like to produce more, but cannot hire additional labour since no more labour is being supplied. This causes the wages and prices to rise in the intermediate term. The rise in wages induces households to supply more labour, so production can rise. Assuming real wages to rise after the demand shock, production will rise less (repressed inflation). Rising real wages stimulate further demand for goods.

When prices start to adjust, both real demand and real wages fall. After full adjustment of the wages and prices the economy returns to the initial Walrasian equilibrium. Due to the assumed structure of investment, the decline in production towards the equilibrium will cause the production to ‘undershoot’ the equilibrium as in the multiplier-accelerator model causing to move the economy in the region of Keynesian unemployment.

Nominal and real rigidities result in quantity-adjustments in the short run. Although no explicit business cycle theory emerged from the new Keynesian approach, it should be possible to combine the accelerator model with the slow price and wage adjustments, as explained above. This shows that to explain the business cycle, using the Keynesian approach, it is not only necessary to assume slow adjustments in prices and wages, but it is as important to assume (and explain) the lagged reaction of investments to changes in production and demand.

3.3 Monetary business cycle theories

The classical monetary theory and the modern monetarists incorporate the quantity approach to money in their models. Given the full utilisation of resources and a constant rate of velocity, the supply of money influences only the general price level. A distortion can have real effects in the short run. In the long run, the economy returns towards its real equilibrium. The following theories see the banking sector as a source of instability, causing nominal and (or) real changes.
Classical monetary business cycles

Hawtrey\(^8\) takes the most extreme position on the monetary theory of the business cycle. He explains the business cycle as a purely monetary phenomenon. The cycle is the result of the incapability of the banking system to adjust the real interest rate when the value of money changes. If the economy is on a long-term growth-path, real demand for goods equals real supply. Depending on the (expected) sales and the interest rate, firms decide on the desired level of stocks and adjust their actual stocks.

Consumers divide their income over consumption and desired cash balances. Hawtrey assumes a lag between a rise in income and the moment at which the cash balances become deposits. A rise in deposits increases cash reserves of the banks. The banks supply interest-bearing credits to the firms to finance their production. In the expansion phase of the business cycle the banks have large reserves at their disposal. This stimulates the supply of credit at low interest rates. The availability of cheap credits causes a rise in production, increasing income. The rise in income raises expected sales, causing the desired level of stocks to rise also. Prices increase when the demand for goods increases, so the rise in profits stimulates (speculative) stockholdings. This stimulates the demand for credit and the expansion of production. Wages adjust slowly as production increases. The lag between the rise in income and the rise in bank reserves causes demand for credit to rise faster than reserves. The ratio of credit to reserves also rises. In reaction, the banks increase the interest rate and lower credit supply. The desired level of stocks declines because of the higher financing costs. The credit rationing and the decline in desired stocks decrease production. Depression will result. During the depression less credit is demanded but initially deposits still rise, as income is turned into deposits. The ratio credit-reserves declines toward its desired level. When credits are at or below their desired level, banks start to lend again and the interest rate declines. This creation of money through bank lending causes the economy to expand. There can be transitory real effects due to the monetary cycle but there are no structural changes.

A combination of the pure monetary cycle and the real theories of the former section was developed by a group of economists, especially Wicksell and Hayek. A common feature of these kinds of business cycle theories is the view of the monetary sector as cause of the cycle, followed by changes in the structure of production, which cause the fluctuations to persist. We present the cycle as described by Hayek (see Hayek(1931, 1933), also see Haberler(1939), Kyun(1988) and Zijp(1992) as examples of this approach).

\(^8\)These ideas are older, also see S. Hollander, Mill on growth, the cycle and the law of markets, in Eatwell(1987, 474-475), quoting Mill, J.S., The currency question, Westminster review, 41, June 1844, 579-598. Mullineux(1990, 64), quotes Mill, J.S, Principles of Political economy, 1848. Haberler(1939, 15), notes: "many of the propositions advanced by Mr. Hawtrey [...] have had a long history and were given an early expression in A. Marshall’s evidence before the Gold and Silver Commission, 1887". Haberler and Mullineux list Hawtrey’s work, starting with Good and bad trade (1913) to Capital and employment (1937).
**The monetary sector**

As in Hawtrey’s theory, households offer their savings to the banks. The banks offer funds on the money market (loanable funds), which can deviate from the savings. Following Wicksell, Hayek distinguished between the natural interest rate and the market interest rate. The natural rate is the rate at which real savings equal real investments. The interest rate at which supply of loanable funds and demand for credits are equal, is the market interest rate. Demand for loanable funds (investment) is determined by the market rate of interest. The supply of funds depends on the savings and bank credit. When the investments are above savings, nominal demand exceeds production, so prices rise.

Suppose the market rate to be below the natural rate. Investment projects that were not profitable at the natural interest rate, are now carried out. The demand for loanable funds rises above the amount of real savings, so credits are supplied (supply of and demand for loanable funds are in equilibrium). Nominal demand for goods becomes larger than supply, inducing nominal prices to rise. A rise in prices lowers the real interest rate and increases expectations of inflation, stimulating speculative production. Both tendencies increase the demand for credits further. As before, this boom would end when investment opportunities decline. A rise in the credits/reserves ratio and a decline in profitability cause the banks to restrict their unsecured loans. This creates a recession as investment declines further. These changes are largely monetary in nature. The interest rate, however, also has a function in the real sector.

**The real sector**

Hayek describes the production of goods as a succession of stages of production. The production process has the final purpose of producing consumer goods, but between the purchase of raw material and the sale of the consumption good, the raw material is transformed using the factors of production, each stage using a different method of production. The factors of production are divided over the production process as a whole, based on the relative factor prices and profitability. The capital intensity of the method of production declines towards the production of the final stages: closer to the final sales, capital is used in a lower proportion of production inputs. The distribution of capital over the stages is determined by the interest rate paid on the funds loaned to finance the initial investment: when the interest rate (the capital costs) increases, capital flows towards the production of the less capital-intensive production of consumption goods.

A monetary expansion lowers the market rate of interest, lowering the costs of capital. The profitability of the production of intermediary goods

---

9For convenience we assume every stage of the transformation process to take place in a different firm.

The Business Cycle
(including capital goods) rises. Firms producing these kinds of goods increase their investments: demand in the capital-intensive sector increases, inducing a rise in the production and prices of capital goods. Both the inflation and the additional demand stimulate investment, and so future supply. As long as the additional money stays within the intermediary sector, expansion continues.

As growth in the intermediary sector increases, the demand for labour in this sector rises. The labour supply is assumed constant, so wages in this sector have to rise to attract labour from the consumption goods producing sector towards the intermediary sector.

Demand in the consumption goods producing sector remains initially equal. The profitability in the consumer goods sector declines because of the rise in the price of capital and the rise in wages. To defend profits, firms increase their prices. As wages rise, nominal demand increases too, but real demand of consumer goods declines (forced savings). However, total production increases because of the investment boom (demand for capital goods).

*Monetary and real causes of the crisis*

The expansion ends because of monetary and real problems. As seen above, the rise in demand for loanable funds during the expansion increases the credit/reserve ratio. The propensity to lend declines. There will be a tendency to increase the market interest rate.

In the real sphere, disproportionalities appear. The demand for consumption goods has risen as the wages rise. Producers of consumer goods would like to expand production, but the price for credits has risen, whereas the price of intermediary goods does not decline because of the higher costs of capital, so the consumption market exhibits excess demand. The price of consumer goods rises more than proportionally. Because of the former investments, capacity in the intermediary sector has risen, but due to the rise in the market rate and the restriction on credits, demand declines: the intermediary market displays excess supply.

Real savings decline as more income is spent on consumer goods, to compensate for the forced savings, again stimulating inflation.

This deadlock of excess supply of intermediary goods and excess demand for consumer goods lowers the propensity to invest further. The market interest rate rises above the natural rate and depression sets in. The depression moves the economy back towards its old proportions, as given by the equality between saving and investment. As Hayek remarks, however, the redistribution of funds over the stages of production involves destruction of real capital and unemployment: this restructuring of the economy has severe real effects: investment projects become unprofitable and firms go bankrupt. As the market rate moves towards the natural rate,
banks restore their reserves and profits in both sectors return to their old level. The economy returns to its equilibrium position. The accumulation of reserves by the banks continues up to a level at which the banks discover they have more liquid funds than required, given the desired reserve ratio. The market rate is again adjusted downward, below the natural rate. Profitability on capital rises, investment increases again, so demand for credits also increases. A new expansion is on its way, with the accompanying lengthening of the production process.

Comparing Hayek's theory to that of Hawtrey, both theories see the primary cause of the business cycle in the behaviour of the banks. The credit creation of the banks weakens the linkage between real investment and real saving. Banks can supply credits above or below the amount of deposits by increasing or decreasing their reserves. Creation of money stimulates investment both directly and through a lowering of the interest rate. According to Hawtrey, the real effects of this expansion are only transitory and the mode of production is not influenced. The main changes occur in the production for stocks (with temporary effects on employment). When the supply of credit returns to its normal level, production and employment also return to their equilibrium values. In Hayek's opinion, those monetary adjustments influence the production structure because of the role of the interest rate in the distribution of capital over the different stages of production. Disturbances in the monetary sphere have significant effects in the real sphere. The controversy between Hayek and Hawtrey concerns the duration of the effect of monetary disequilibrium on the real sector. The real theory on the business cycle, as presented in 3.2, does not reject the influence of the monetary sector, but sees it as secondary to the real disturbances. Even in the absence of monetary distortion the economy would fluctuate. In Hayek's view fluctuations cease to exist, when the natural rate equals the market rate. Yet, there is considerable agreement on the effect on production and employment before the economy returns to its 'normal' proportions.

**Modern monetary business cycles**

In modern monetary economics, Greenwald and Stiglitz (1993) develop a model in which the monetary sector restricts the growth of the real sector because of imperfect information and risk aversion of the banks. The banks have no instruments to assess the risks of the investment proposals, which firms want to finance through bank loans. Secondly, banks try to avoid the costs associated with bankruptcy, behaving in a risk adverse way. The level of lending (quantity) and the interest rate (price) are set depending on the overall expectations and the net worth of firms. Macroeconomic variables influence the expectations and the net worth of the firms, determining part of banking behaviour. The banks influence the aggregate economy through their interest rate and loan policy.

Suarez and Sussman (1997) present a variation of the Stiglitz-Weiss economy, in which "the only imperfection is that an entrepreneurial
decision which affects the returns from investment is not verifiable and hence subject to moral hazard” (Suarez and Sussman(1997), i). Taking a three-period overlapping generations model, in which there is a probability of default in the third period, Suarez and Sussman show the existence of a cobweb cycle of period 2 in prices.

Another approach, which emphasises the role of financial factors in business fluctuations, is the ‘Fisher-Kindleberger-Minsky’ approach (Gatti(1992), 134). The financial instability hypothesis, as it is also called, combines behaviour of banks as innovative entrepreneurs with the investment behaviour of innovative firms. Innovations earn (temporary) monopoly rents for those firms that can capture market power. “Banks are also entrepreneurial firms that innovate to earn profits. As financial institutions innovate, the financial system evolves, finding new ways to finance the profit-seeking behavior of other institutions” (Wray(1992), 163).

Expected future demand and profits, comparing the supply price of capital with its demand price, determine investment. Financial factors are part of the supply price because of the financing of the investment good, but also of the demand prices, because of the discounting of future expected profits. Business fluctuations originate because “financial markets may operate to reward speculation-leading to a speculative boom. A crisis can occur if finance costs rise, if liquidity preference rises, or if income flows turn out to be less than expected. Endogenous processes tend to ensure that one of these (or all three) will, in fact occur”(Wray, op cit., 167-8, also see Minsky(1994)).

Minsky’s financial instability hypothesis gave rise to a wide line of empirical and theoretical research. Some of the elements of the financial instability hypothesis, for example the role of banks as profit-seekers, will return in the following chapters.

4 The political business cycle

The approaches above explain the cycle according to macroeconomic relationships. A different approach to the business cycle is taken by the public choice approach. This field of research aims to model the influence of political groups as politicians, government officials and lobby-groups. Economic fluctuations are explained as the result of actions taken by these groups, aiming at political influence or power.

A simple model of political behaviour, based on Nordhaus(1975), can illustrate the way in which those groups can cause the business cycle. Nordhaus uses the Phillips curve: unemployment affects money wages and money wages affect prices. In the short-term, there is a trade-off between inflation and unemployment, but in the long run the natural rate of unemployment hypothesis is assumed. Individuals maximise a utility-function, preferring combinations of non-fluctuating prices and low
unemployment to combinations of inflation and high unemployment. These individuals are assumed to be rational in their preferences, but ignorant of the trade-off between inflation and unemployment.

There are two parties, the government (the ruling party) and an opposition party. The individuals, the voters, make a decision which party to vote on, based on the expected behaviour of the political parties (determined by adaptive expectations) and the realised outcome on unemployment and inflation. The policy makers are assumed to maximise the chance of being re-elected, knowing the preferences of the voters. Nordhaus(1975) shows that the democratic process results in a situation with lower unemployment and higher inflation (ignoring the preferences of future generations) compared to the optimal long run equilibrium (Nordhaus(1975), 178-179). Nordhaus starts his analyse of the political business cycle by pointing out a difference with Kalecki’s political cycle. Kalecki argues that the political cycle has its roots in the unequal distribution of political power between ‘captains of industry’ and the ‘workers’. A stimulating policy enlarges employment, and thus wage-demands and inflation. The economy will be depressed deliberately, as an instrument in the class struggle.

Nordhaus(1975) assumes a democratically elected government. Election periods are constant and the voters have a decaying memory: the more recent the experiences, the higher their weight in the memory. As he remarks: “[...] if voters can predict future events perfectly and discount the future at the social discount rate [...] the system will tend toward the social optimum [...]” (Nordhaus(1975), 182, note 2). This gives the policy makers the opportunity to manipulate the economy and the voters. Shortly after election, they fight inflation by allowing unemployment levels to increase. When election day approaches, they will turn to a stimulating policy, raising inflation and depressing unemployment. “The optimal partisan policy will lead to a political business cycle, with unemployment and deflation in early years followed by an inflationary boom as elections approach” (Nordhaus(1975), 185). Confronting his theory with the data for different countries Nordhaus(1975) finds it fits the data well10. Especially as the data for those periods (for the United States) that did not fit the theory, the governing party lost the elections. This brings Nordhaus to advocate for a system of indicative planning, to broaden the need for consensus over the policy and reduce the possibilities for ‘egoistic’ policies. The Dutch ‘poldermodel’ can be seen as a variation of Nordhaus’ indicative planning, in which the negative effects of partisan policies are internalised by the parties.

*Other political business cycle theories*

Saint Paul(1994) has focused on the explanation of the high level and persistence of European unemployment. He argues that rigidities in the

---

10 The exception being the United Kingdom. This is caused by the balance of payment problems during the early sixties.
labour market, which obstruct market-clearing, are not motivated by social reasons, but, because they benefit the employed (by raising their bargaining power). This argument can be used to extend Nordhaus' political cycle by introducing two additional factors:
1. The influence of employed on policy makers is larger than the influence of unemployed, and
2. the employed care more about inflation (real income), whereas the unemployed care about employment, so:
"since the employed, even at high unemployment rates, are much more numerous than the unemployed, it is evident that policy is determined by the interests of the employed rather than those of the unemployed" (Saint-Paul(1994), 3). In this case, the party in government rather raises rigidities in the labour market (and unemployment) than fund a buoyant policy (followed by tax rises or inflation), to get re-elected.

Beetsma and van der Ploeg (1992) also explain the cyclical occurrence of inflation as the outcome of the democratic decision process\(^1\). They assume a link between inflation and inequality. The government chooses a policy that maximises the utility of the median voter. In an economy with large inequality, the median voter will be in the lower income classes. Policy is aimed at a rise in effective demand (running a large deficit with low tax rates). This rise in demand increases imports and uses up idle production capacity. The imports will deplete foreign reserves, and domestic prices increase, until, eventually, a capital flight and expectations of exchange rate adjustments, force the government to lower the budget deficit and thus inflation\(^2\). These policies are supported by the lower income classes when the inflation affects their consumption possibilities, yet the deflationary policies could lower living standards below the initial situation. After time the 'populist policy cycle', as Beetsma and van der Ploeg label this process, could take off again.

The ambition of the politicians for re-election, which provides a continuous impulse in these kind of models, is exogenous to the economic system. The occurrence of the economic cycle, however, depends on the economic structure. A pre-election rise in government spending increases employment, but this increase is assumed to be only temporary for a number of reasons (inflation, structural unemployment). The post-elections depression is inevitable\(^3\). As the behaviour of the banks before, the behaviour of the politicians results in endogenous fluctuations.

\(^1\)The analysis in Beetsma and van der Ploeg(1992, 1) is directed at Latin American countries, but could be applied to all countries in which "inequality and social unrest sow the seeds for populist policies of redistribution and expansion".
\(^2\)These deflationary policies were enforced by the stabilisation programme of the IMF, Beetsma and van der Ploeg(1992, 6).
5 The business cycle and growth

Several business cycle theories assume the economy to grow. These approaches use the Solow-residual as impulse (real business cycle) or assume a 'natural' growth path (growth cycles; predator-prey models). Growth causes fluctuations either because of the adjustment process along the growth path or because the growth rates of different variables diverge, causing pressure on the system. They, however, do not provide an explanation for the growth path itself.

Aghion and Howitt(1998) develop an Schumpeterian endogenous growth-model. Growth results because of innovations, and involves creative destruction: new inventions make old techniques obsolete. The occurrence of new innovations depends on the amount of labour that is used in research. Growth results when the positive spillovers of the innovation (learning effects) are larger than the negative spillovers (creative destruction).

If growth is endogenous, there are two approaches towards cycles: cycles are caused by growth (Kondratieff cycles) or, alternatively, the business cycle causes the growth rate to fluctuate.

Cyclical growth
Growth can take the form of cycles as there is a negative relationship between future and current research activities: more future research (a higher chance of a future innovation) lessens the future profit opportunities of present innovations (creative destruction because of the future innovation). This mechanism can result in a perfect foresight equilibrium (PFE) of cyclical growth (Aghion and Howitt(1998), 63/64)14.

Another reason for cyclical growth is that recessions, as Schumpeter stated, pave the way "for new achievement of the kind which has created civilization and made the greatness of this country" (cited in Stiglitz(1993), 3). Aghion and Howitt(1998) formalise this concept by introducing general purpose technologies (GPT's, i.e. the steam engine, computers). The spread of technological knowledge is modelled by usage of an epidemiological model.

There are three stages: discovery of the GPT (stage 0), experimentation (stage 1) and implementation (stage 2). After the discovery of a new GPT labour is devoted to research, to determine the usage of this new technology in the own sector or firm. As more firms assign workers to experiment with the

---

13If the economy has Keynesian characteristics, the government deficit generates new income, so the rise in taxes or inflation is not evident. In this case the mechanism of Nordhaus generates "political" growth.
14Aghion and Howitt(1998) describe this PFE as: \( n_t = \psi(n_{t+1}) \), with a two-cycle as \( (n^0, n^1) \) such that \( n^0 = \psi(n^1) \) and \( n^1 = \psi(n^0) \), \( n \) is the amount of labour in research.
new technology, less labour is available for productive activities (using the old GPT). Output declines further as more firms enter stage 1. The implementation of the new GPT depends on own discoveries of applications within the own firm or on imitation of other firms within the same sector (Aghion and Howitt(1998), 254).

Imitation becomes easier when more firms make the transition from stage 1 to 2. This accelerates the overall transition. Aghion and Howitt(1998) estimate the transition to be at its peak at (about) 20 years after the discovery of the new GPT. Macroeconomic production rises as labour is redirected towards the new productive activities (using the new GPT): “cyclical downturns may be the price that society needs to pay in order to implement the GPT’s that deliver long-run growth” (Aghion and Howitt(1998), 244).

Growth based on the cost of R&D is also central to the ‘extended cost-disease model’ of Baumol and Wolff(1992). They explain fluctuations in the growth rate from the effects of growth itself. The economy is divided in two sectors. Information is produced in one sector, using labour as an input (R&D). The other sector produces goods, using both information and labour as inputs. The labour productivity depends on the amount of information used in the production process. Demand and supply of information are determined by the price of information, relative to the price of the good. Growth (a high demand for R&D) causes the price of information to rise. This depresses demand for R&D in the next period: labour productivity declines and a depression sets in. A cobweb cycle in prices and demand in R&D follows. Depending on the way, the interaction between sectors is modelled, different kinds of behaviour (including chaotic fluctuations) can be observed.

**Business cycles to growth**
If growth is generated endogenously, short-term changes influence long-term growth. Question is then if growth is a virtue of bad times (Aghion and Howitt(1998)), or a characteristic of good times?

**Growth through good times**
Changes in the level of production have an impact on growth through the ‘learning by doing’-mechanism or by stimulating research. Suppose the money stock to rise. Because of nominal rigidities (see the new-Keynesian approach above) or informational rigidities (see Lucas’ rational expectations) production rises in the short run. In the regular neo-classical tradition, prices rise as menus are adjusted and information is spread throughout the economy. In the long run the money stock-price ratio will return towards its old level, as production does too.

When the growth rate is determined by learning-by-doing a temporal rise in employment causes knowledge to grow, so the growth rate increases permanently. If the amount of labour spent on research depends positively
on the level of production and/or negatively on the wage rate (assuming the real wage to lag behind the rise in production, so real wages decline as prices rise), labour is redirected from (direct) productive usage towards R&D. The rise in R&D-efforts raises the chances of the appearance of new innovations.

Both mechanisms assume a positive relationship between growth and the period of expansion. Stiglitz(1993) addresses the same relationship, using a model of technical change with credit rationing. The pressure on capacity and the availability of resources during the expansion phase cause innovations to develop. There is no incentive to modernise production when the utilisation rate is below its normal level. "Typically, firms also reduce their expenditure on R&D and productivity enhancing expenditures. The reduction in output reduces opportunities to "learn by doing" (Stiglitz(1993), 3). Technical progress, in his opinion, takes place because of the high level of costs and the pressure on the production capacity in a boom. This leads to growth, but eventually to a situation of overproduction: the cause of a recession. Innovation takes the form of labour saving new techniques. This increases unemployment levels (given a low state of demand). The instalment of new techniques has to be outweighed by the rise in production otherwise it will enhance the recession, instead of leading to a boom.

When the rise in production favours growth, productivity should be procyclical. This procyclical behaviour (especially of the Solow-residual) is also found in empirical data. Recent research, however, finds evidence of countercyclical behaviour of productivity (Aghion and Howitt(1998), 239).

Growth and the virtue of bad times
Aghion and Howitt(1998, 239) follow Schumpeter in describing a recession as a "cleansing mechanism for reducing (or eliminating) organizational inefficiencies and resource misallocations" (also see Hayek's theory on the business cycle above). There are several reasons:
1. the 'lame duck' effect: a recession forces inefficient firms out of business, but also provides higher entry costs for efficient firms;
2. the opportunity costs of (non-productive) activities as reorganisations are low during a recession;
3. the 'disciplinary' effect: when the majority of firms reorganise during a recession, those who do not will go bankrupt.

Aghion and Saint Paul(1991) modernise Schumpeter's vision in the 'opportunity costs-approach' of growth. Short-term recessions stimulate overall growth because the opportunity costs (foregone profits) of investing in technological improvements and reorganisations are low in a recession compared to the opportunity costs of those activities during a boom. There are two important differences between the opportunity-costs approach and the Schumpeterian (Austrian) theory. The starting point of the latter is a
theory of the business cycle, involving the actual destruction of existing capital, through the ending of investment projects and bankruptcy. Aghion and Saint Paul rely on an exogenously generated business cycle, ignoring these effects of the downturn.

Schumpeter related growth to the great inventions, the link between growth and fluctuations is the creative entrepreneur who carries out these new inventions or rearranged factor inputs in a more efficient way, involving risk taking and profit seeking. Aghion and Saint Paul also stress the importance of new ideas and technical progress. In their model, however, innovation is realised through expenditures on R&D and implemented automatically without reference to the (time-consuming) process between invention and innovation.

Van Ewijk(1994) extended the opportunity costs-model of Aghion and Saint-Paul. Following Aghion and Saint-Paul(1991), the business cycle is represented by exogenous fluctuations of the profit rate. Firms are assumed to make two decisions. First, they have to decide if, given the state of the profit rate, they will be active (enter the market place) or become inactive (exit the market place). The costs of becoming or remaining inactive are twofold. There are entry-costs when the firm decides to become active again and the knowledge (labour productivity) deteriorates because of the absence of ‘learning-by-doing’. If the firm is active, the next decision to be made regards the amount of labour dedicated to (non-productive) research, determining the growth rate of (labour) productivity. Using several restrictions and assuming profit maximisation, van Ewijk(1994, 19) concludes: “recessions, in particular deep and long recessions, lead to inactivity and idle resources, they have a negative effect on learning-by-doing and cause ‘rusting’ of existing knowledge [but] on the other hand, mild fluctuations which remain within the ‘full employment’ corridor may be beneficial to growth..[.....]..this may help to explain why long and smooth cycles like Schumpeter’s ‘Kondratieff wave’ involve a positive relationship between cycles and growth, while short and vehement ‘Keynesian cycles’ tend to be detrimental to long-run growth”. Aghion and Howitt(1998) conclude, based on empirical research, that small recessions are beneficial to growth, whereas large fluctuations are obstacles for growth as they “increase the option value of delaying irreversible (R&D) investments” and “delay the diffusion of technological knowledge” (Aghion and Howitt(1998) 243).

Helpman(1992), focusing on international growth differences, distinguished between two sorts of innovations: product- and process-innovation. Although Helpman (1992) did not consider fluctuations, this distinction can be applied to the process of innovation over the business cycle. During a depression, in a declining market and a low wage rate, is R&D concentrated on product innovation, leaving the labour-intensity of the production process unchanged. When the economy recovers and labour costs rise, innovation is turned towards process-innovation, lowering production costs. In this interpretation of Helpman’s distinction growth is advanced by
the occurrence of short-term fluctuations, either by raising quality of by lowering costs of production. Matsuyama(1999), for example, models an economy in which there is a ‘Solow regime’, exhibiting classical growth through factor accumulation and a ‘Romer regime’, in which growth is innovation led.

6 Conclusion and summary

This chapter gave a survey of several theories that are brought forward to explain the short-term fluctuations, which are classified as ‘the business cycle’. They can be classified using Frisch(1933)’s impulse-propagation problem. At one end of the spectrum are the ‘perpetuum mobile’-theories. Their research is aimed at an explanation of the mechanism responsible for the fluctuations, the propagation mechanism (paragraph three). Contrary to those theories, there are ‘Deus ex machina’-theories, assuming an inherently non-fluctuating economy, at least in the long run. The cycles are seen as adjustments after an exogenous shock, the impulse (paragraph two).

Frisch(1933, 197) asks “[H]ave these dynamic laws deduced from theory and showing damped oscillations no value in explaining the real phenomena ..?”. He concludes that exogenous impulses are necessary to explain persistent fluctuations. Contrary to this, Haberler(1939, 10) states that “[F]or various reasons, it seems desirable, in the explanation of the business cycle, to attach as little importance to the influence of external disturbances.. large swings in the direction of prosperity and depression as we find them in real life are difficult to explain solely by exogenous forces; and this difficulty becomes an impossibility when the alleged “disturbances” do not themselves show a wavelike movement.. historical experience seems to demonstrate that the cyclical movement has a strong tendency to persist, even where there are no outstanding extraneous influences at work which can plausibly be held responsible”.

Economists such as Schumpeter and Friedman accept the fact that disturbances are both real and monetary. The different ways of propagation give rise to endogenous persistent fluctuations:
- monetary mechanisms: involving the behaviour of banks, either only affecting prices or also affecting production;
- real mechanisms: the investment behaviour and the distribution of income;
- psychological mechanisms: self fulfilling and counteracting behaviour causes fluctuations;
- exogenous mechanisms: non-economic behaviour can cause economic fluctuations (political cycles).

15It has to be noted that stability of the economy in this sense does not necessarily mean full employment. As is shown by the simple multiplier-accelerator model, a long term equilibrium can exist with unemployment.
There is some common ground. The different endogenous theories can be combined into a general theory on the business cycle. Once it is accepted that the interest rate plays a role in determining intertemporal choices, there is an interaction between the real and monetary sphere. This relationship is enhanced when the psychological state of mind of businessmen and bankers (Mills’ psychological cycles) is taken into account. Take for example, a psychological boom, caused by a (electorally motivated) decline in taxes (Nordhaus’ political cycle). Both entrepreneurs and bankers expect this to raise production and increase profitability. Investment is increased and the availability of credits expands. In the short run production rises more than the initial tax reduction because of the additional (credit-financed) demand for investment goods (Hawtrey-Hayek’s monetary cycle). Assuming the taxes to be equally reduced for producers and consumers, both the demand for producer goods as the demand for consumer goods has risen. The increased credit-financed demand for producer goods competes with the additional demand for consumer goods (Haberler’s disproportionality argument). The price of consumer goods rises relatively to that of producer goods, wages rise (Hayek’s transition of labour, Goodwin’s conflict over the income distribution) and savings decline (countering forced savings). When the past investments are implemented, production capacity in the producer goods sector rises (Kalecki’s investment cycle).

When real savings diminish and there is an excess capacity in the producers sector, credits will be constrained (Hawtrey’s credit/savings ratio). This creates a pessimistic state of mind, causing a depression to set in. If elections are held in the foreseen future, the depression damages the chances for re-election, creating an incentive for politicians to lower taxes further. The duration of the fluctuations, caused by the combination of propagation mechanisms, depends on the way the different groups (workers, entrepreneurs, bankers and politicians) interact.

A macroeconomic model of the business cycle, which would capture the common features as described above, has to take into account:
- the relationship between the interest rate, savings/credits and the behaviour of the banks;
- the conflict over income (profits versus wages) between workers and entrepreneurs;
- the decisions on investment and production, within firms, based on prices and costs;
- behaviour of non-economic groups, such as politicians, who have a stake in the performance of the economy and the possibilities to influence the behaviour of the other stakeholders.

Growth remains implicit in many of the business cycle theories. Growth itself can be a source of long-term cycles as the amount of labour devoted to research and development fluctuates. Growth can also provide an exogenous stimulus to short-term fluctuations. Uneven sectoral growth
rates can cause pressure on supply (underconsumption theories: former production is present income) or on the labour market (Goodwin cycles: the accumulation of capital deviates from the growth of the working force). Other theories (RBC-models) explain short run cycles as part of the adjustment process when innovations increase the growth rate. When the growth rate is endogenous, short-term fluctuations influences the long-term growth rate. The question is which situation, depression or boom, causes growth to accelerate? In both situations, small fluctuations seem more stimulating then large cycles.

The research on the relationship between growth and the business cycle is relatively young and conclusive deductions are not yet made concerning the exact relationship between growth and cycles.

Reasons for government interventions are threefold:
- to stabilise the economy as the cyclical adjustments towards a new equilibrium take too much time (Lucas’ rational expectations, RBC-cycles);
- to stabilise the economy because the structure itself is such that a new equilibrium will not be reached;
- to prevent the occurrence of depressions by lowering excess growth (Hayek). This assumes both sincere behaviour from the government (Nordhaus’ political manipulation of the economy) and good knowledge of the economic structure (to determine the goals and ends of the intervention).

Yet, another problem remains to be solved. Returning to the impulse-propagation problem, Goodwin(1990, 63) can be paraphrased: when Tinbergen (see for example Tinbergen(1943)) and Frisch introduced (dynamical) mathematical methods in economic theory, they introduced a new consistent way of looking at the economic system, but they also destroyed a large body of theoretical knowledge. Many mathematical dynamic systems could and can not (yet) be analytically solved. This restricts the mathematical representation of an economic theory to certain forms. These systems can generate fluctuations, but within severe restrictions. Even the so-called chaos and catastrophe theory (for an expose on catastrophe theory in economics, see Rosser(1991)) requires certain boundaries for the system to be solvable.

It is obvious that Frisch(1933), at the time and given the state of mathematics, concluded that economic systems are, generally, stable and fluctuations dampened. This denies the idea of an endogenous economic cycle. Yet, successive periods of high and low growth can be seen in the economic data. That economic fluctuations are difficult to model and that business cycle-models are difficult to analyse in a general way, does not deny the importance of analysing economic fluctuations.

---

16See for example Frisch’s comment on Aftalion: “I do not think that Aftalion’s analysis as originally presented by himself can be characterized as a determinate analyse. By putting his arguments in equations one will find that he does not have as many equations as unknowns”, Frisch(1933, 181).