Overdeterminacy and endogenous cycles: Trygve Haavelmo’s business cycle model *

André K. Anundsen        Tord S. Krogh
Ragnar Nymoen            Jon Vislie†

19 February 2014

Abstract
The econometrician Trygve Haavelmo pursued a research program in macroeconomic theory that was highly original for its time. We present his macro model for an economy with deregulated financial markets and a policy determined interest rate path. Disequilibria arise in the interface between asset markets and the real economy. A mismatch between the marginal return to capital and investors’ required rate generates endogenous switching between recession and full employment regimes. Haavelmo regarded the ‘switching mechanism’ as a substitute for liquidity constraints, and together with his ideas of price dynamics, there is a clear Keynesian and Wicksellian influence on his macroeconomic theorising.

Keywords: Investments, Business Cycles, Monetary policy.

JEL classification: E22, E32, E44, E52

1 Introduction
Trygve Haavelmo (1911-1999) is well known for his seminal contributions to econometric methodology, and in particular “The Probability Approach in Econometrics” (Haavelmo, 1944). What is less known, is that Haavelmo also devoted much effort to developing macroeconomic and monetary theory, as well as models for business cycle dynamics and secular trends. His results were original, yet strongly influenced by Ragnar Frisch (his teacher), John M. Keynes and Knut Wicksell.

*We would like to express our gratitude to the editor Neri Salvadori and to two anonymous referees, for constructive critique and encouraging comments. Thanks also to Olav Bjerkholt, Sheetal K. Chand, Eilev S. Jansen, and Asbjørn Rødseth for valuable comments. Krogh is grateful for funding provided by the Norwegian Research Council, project number 195135. Nymoen and Vislie are affiliated with the Department of Economics, University of Oslo. Krogh is affiliated with Statistics Norway, and Anundsen is affiliated with Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank.

†Corresponding author: Jon Vislie, Po BOX 1095 Blindern, N-0317 Oslo, jon.vislie@econ.uio.no.
The theoretical approach of Haavelmo’s research after WW-II can be seen as a natural response to his own assessment of how econometrics had developed over a quarter of a century: statistical methods had been improved with great pace, while there was less progress in the development of economic theory. This meant that it was a greater need for better theoretical models than further development of statistical techniques.¹

Haavelmo’s interest in macroeconomic fluctuations was rooted both in his own personal experience from the 1930s, the inspiration he got from Frisch, and the early research he undertook prior to revolutionising econometrics. As we document in this paper, monetary and business cycle theories continued to interest him throughout his career. Already at the New York meeting of Econometric Society in December 1941, in the intense period when “The Probability Approach” was coming to a completion, he outlined the main ideas of a paper entitled “Wicksell’s Theory of Interest Rates and Prices” (Haavelmo, 1942), and he returned to that theme several times, both in his 1960 treatise on investment theory and in several lecture notes in Norwegian.² Our presentation of his core ideas are therefore important to get a more complete picture of Haavelmo’s scientific production.

Upon finishing high school, Trygve Haavelmo understood that, because of the ongoing depression in the 1930s, it was impossible to find employment directly. He therefore decided to become a student of economics at the University of Oslo.³ From the autumn of 1930 to the spring of 1933, Haavelmo studied economics in Oslo, with Ragnar Frisch as his most influential teacher. After graduation, he was employed as a research assistant, and later as the “Chief computer” at the Institute of Economics, a position he retained until August 1938. In a letter of recommendation, written by Frisch upon Haavelmo’s leave from the institute, Frisch wrote that “part of our work have been directed towards the analysis of time series with a particular view on the determination of cycles”.⁴ Hence, Haavelmo was involved in macroeconomic research even before he started on the “Probability Approach”.

The work Frisch referred to was part of his more than decade-long research program in macro dynamics and business cycle modelling (Bjerkholt and Lie, 2003). Haavelmo must have been introduced to Frisch’s program as a student, and he took part in the research when he became Frisch’s assistant. It was therefore natural that Haavelmo’s European journeyman year from December 1937 to June 1938 was a mission with two aims: to study mathematical statistics and to visit the leading European business cycles researchers in Berlin and Geneva (Tinbergen in particular).⁵ Haavelmo had also lived in London for

¹In his Presidential address to the Econometric Society, Haavelmo stated: “There are already some econometricians who are engaged in work on the fundamentals of economic theory along the lines I have indicated. To them we should give all possible encouragement.” (Haavelmo, 1958b, page 357).

²At the New York Meeting, Haavelmo discussed Wicksell’s cumulative process in a dynamic model where savings were equal to investments and where the money market rate of interest was controlled by “the banks”.

³Interview with Trygve Haavelmo by the Norwegian Broadcasting Cooperation (NRK) in connection with the Nobel Prize in 1989.

⁴Letters dated December 1st 1937, from Professor Ragnar Frisch to Professor Ernst Wagemann at the Institut für Konjunkturforschung, and to Karl Stumpff who directed the Meteorologisches Institut’s department for ‘Periodenforschung’. File labeled “Korrespondanse med Trygve Haavelmo”, Box 152.A, in the archive of the Department of Economics, University of Oslo.

⁵Haavelmo business cycle paper in Econometrica (Haavelmo, 1940) acknowledged that he was building on unpublished results by Frisch about methods for determining principal characteristics of cumulative
three months in the autumn of 1936. He attended the lectures in mathematical statistics
given by Professors Neyman and Pearson at the London University College. During the
stay he also took part in Professor Hayek’s business cycle seminar at the London School of
Economics. Haavelmo joined up with Frisch at the meeting of the Econometric Society
held in Oxford that year. At this conference, the discussion of the General Theory was
an important item on the agenda, see Bjerkholt and Lie (2003). Interestingly, in a letter
to Jacob Marschak, dated November 15 1944, in connection with the balanced budget
multiplier paper, Haavelmo wrote that “this paper is a fragment, torn loose from quite
a volume on Keynesian economics that I have managed to piece together during the last
two or three years. I have thought of using it as a textbook for the students in Norway,
if and when I get a chance.”

Against this background, it is not surprising that Haavelmo continued to work on
macroeconomics later in his career. Boianovsky (2002) shows that Haavelmo must have
discussed theoretical issues and macroeconomic models with Patinkin at the Cowles Com-
mision and that there are clear traits of this influence in his seminal offering on investment
theory (Haavelmo, 1960).

Central to Haavelmo’s macro theories was the joint modelling of financial markets
and of the real economy. This approach was inspired by Knut Wicksell (1898, 1906), an
influence that began already when Haavelmo was a student. His analysis of monetary
theory can in fact be seen as complementary to the arguments underlying the cumulative
process.

By taking into account the dual role of capital, both as a factor of production and
as a financial asset, Haavelmo demonstrated how a fundamental overdeterminacy will
arise if the monetary authorities choose to use the rate of interest as a policy instrument.
This is due to the conflicting “signals” sent between the financial markets and the real
economy regarding the equilibrium return to capital. In the real economy, a given level
of aggregate demand determines employment and thus also the real rate of return to
capital. This is the rate of interest that clears the rental market for capital. On the other
hand, equilibrium in the financial market requires that the various rates of return in the
economy satisfy what Haavelmo dubbed the law of indifference in the capital market. For
a fixed money market interest rate, this law defines investors’ required return to capital.
For all markets to be simultaneously in equilibrium, the actual rate of return must be
equal to investors’ required rate. But this can only happen by coincidence! Thus, if the
central bank interferes in the financial market by manipulating the interest rate, this
makes the model logically inconsistent or overdetermined.

---

6Letter from Trygve Haavelmo to the University of Oslo, dated January 6, 1937, in file labeled
“Korrespondanse med Trygve Haavelmo”, Box 152.A, in the archive of the Department of Economics,
University of Oslo.

7We owe a thank to Olav Bjerkholt for showing us a copy of this letter, discovered by Bjerkholt in
Jacob Marschak Papers archive, in the preparation of Bjerkholt (2013). It is unclear what Haavelmo had
in mind when he mentioned his idea of preparing a volume, or textbook, on “Keynesian Economics”.
There are no traces of such a volume in any of the papers kept in the archive at the University of Oslo
Oslo. Maybe his offerings on “Keynesian Economics” were integrated in his SMT from 1966, perhaps we
will never know.

8This is a view concisely presented by Haavelmo himself in Chapter 33 of ”A Study in the Theory of
Investment” (Haavelmo, 1960).
By modifying the overdetermined system, Haavelmo formulated a consistent model where investment demand depends on the discrepancy between the actual and the required return to capital, while realised investments are rationed in boom-periods. Integrated into a macroeconomic model, this generates abrupt endogenous cycles with deep recessions and booming periods. Thus, Haavelmo’s model has the inherent feature that macroeconomic cycles arise endogenously from a mismatch between the two rates of return.

In this paper, we present and explain the logic of Haavelmo’s model. Next to Haavelmo (1960), our starting point for understanding Haavelmo’s ideas is the textbook titled “Orientering i makro-økonomisk teori” (Haavelmo, 1966), henceforth SMT. This book was on the reading list for several generations of Norwegian economists, but was unfortunately never translated into English. By paying attention to aspects of Haavelmo’s theories that are relevant for business cycles and monetary policy, our paper supplements the excellent presentation by Moene and Rødseth (1991) of Haavelmo’s contributions to econometrics and economics.

Our account of Haavelmo’s contributions to the study of the business cycle is organised in the following way. The next section defines his key concepts underlying the law of indifference in the capital market, and the fundamental overdeterminacy that arises if the money market interest rate is used as a policy instrument. In Section 3, we show how Haavelmo formulated a theory of investment behaviour where private sector investments are determined by the difference between the rate of return to capital in production and the required rate of return in the asset market. Embedding this investment response in a macroeconomic model, we show in Section 4 how Haavelmo developed a theory of endogenously boom periods and intermittent depressions. In Section 5, we compare Haavelmo’s setup to Goodwin’s well known framework from 1951. We point out that Haavelmo’s model shows traits of influence from both Keynesian and Wicksellian theories, which is noteworthy in itself, given that the theories are often regarded as polar. Section 6 is a concluding discussion about the relevance of the law of indifference for the debate about economic policy within a deregulated financial market, and shows how Haavelmo’s position in that debate was based on the conclusions in his own macroeconomic model.

2 The required rate of return to capital and fundamental overdeterminacy

Haavelmo’s business cycle theory combines neoclassical monetary theory and Keynesian ideas. Haavelmo regarded Knut Wicksell as the champion of neoclassical monetary economics, a view shared by Frisch. Departing from Wicksell’s analysis, Haavelmo’s model accounted for the double nature of money as a medium of exchange and as store of value.English translation of the title: “A Study in Macroeconomic Theory”.

Andvig (1993) also discusses some of these issues. Morgan (1990) is the first academic appraisal of Haavelmo’s Probability Approach to Econometrics, see Bjerkholt (2005, 2007b) for a recent contribution. The reader is referred to Nerlove (1990) for a more critical appreciation of Haavelmo. Haavelmo’s broad influence on the econometrics research program at Cowles Commission in the 1940’s is also documented in Christ (1994) and Bjerkholt (2013).

For Frisch’s view, see Frisch (1952). Haavelmo lectured on Wicksell in 1951, and published a comment about Wicksell’s contributions in Haavelmo (1978).
Likewise, he considered the two-sided role of the interest rate, since it on one hand must reflect the potential for return to investments in the real economy (Wicksell’s natural rate), while on the other hand it must also act as a market clearing price in the money market. The Keynesian nature of the model will become apparent in Section 3, where we discuss Haavelmo’s thoughts on investment behaviour. His position was that, since savings and investment decisions are separate and uncoordinated, the neoclassical position that households’ choose the macroeconomic activity level becomes untenable. As a consequence, these theories are irrelevant as an explanation of secular changes, such as the Great Depression.

The main points of Haavelmo’s analysis are best appreciated once we understand the meaning of two concepts: the required rate of return to capital and the fundamental overdeterminacy. Consider a stripped-down version of Haavelmo’s model of the asset markets, for a closed economy with one private sector and one banking sector. The private sector is regarded as a single group, consisting of households and capital owners, where the latter rent their capital equipment to firms. At some point in time the private sector has a given nominal wealth, $W$, which along with their net borrowing from the banking sector, $L$, is allocated to deposits, $M$, in the banking sector, and to real capital with a nominal value $pK$. $K$ is a physical measure of the stock of real capital, whereas $p$ denotes the price level. At any point in time $K$ is fixed. Suppose that the banking sector as a whole is in equilibrium in the sense that their lending is backed by deposits. Therefore the following accounting relations must be satisfied:

$$W + L = pK + M$$

$$M = L$$

The private sector’s wealth management is derived from its preferences for holding the various assets. Real capital earns a return $r_K$ when being rented to producers. Holding money as deposits in the banking sector yields the return $i$, which is the money market rate of interest. For simplicity, we assume a fixed price level. Therefore, $r_K$ and $i$ can be interpreted as both real and nominal rates of return. The private sector has a preference function defined over profiles of real capital equipment, real money holdings (as deposits) and real borrowing, as given by $V(K, \frac{M}{p}, \frac{L}{p}, \frac{W}{p}, r_K, i, Y)$. The choice variables are $\left(K, \frac{M}{p}, \frac{L}{p}\right)$, while the real value of the wealth, the various rates of return and GDP (as given by $Y$) are treated as exogenous. This way of formalizing wealth management bears resemblance to Patinkin’s approach, see Patinkin (1956). By maximizing the preference function subject to the wealth constraint (1), standard demand functions can be derived; $x = x\left(\frac{W}{p}, r_K, i, Y\right)$, $x = K, m, l$; where $m := M/P, l := L/P$. In a short-run equilibrium, with $Y$ and $p$ taken as given, and with (1) and (2) satisfied, $pK = W$ is implied. We therefore end up with a single equilibrium condition for the private sector’s asset holdings, as given by:

$$m(K, r_K, i, Y) = l(K, r_K, i, Y)$$

The condition can be regarded as a no-arbitrage condition in the short-run, or as the law of indifference in the capital markets; see Haavelmo (1960). This “law” states that
the prices (here: the interest rates) must adjust so that investors are indifferent between their holdings of the various assets.

The condition in (3) implies the following insights: if monetary policy is “passive”, the actual real rate of return to capital, which in the short-run is given by the marginal net productivity of capital in the producing firms, will determine the money market rate of interest. On the other hand, if the money market rate of interest is used as an instrument by the monetary authorities – an active monetary policy – a unique required rate of return to real capital, denoted \( r^* \), will follow from the law of indifference. The required rate of return is defined as the rate of return that capital owners require, for a given money market rate of interest, in order to keep their existing capital stock as part of their portfolio. Haavelmo regarded the derived relationship, under an active monetary policy, as affected by a large number of factors, \( z \), like overall liquidity in the markets, uncertainty etc, in addition to the explanatory factors included in the stripped-down version of his model. Hence, in a world where the money market rate of interest is set by the government, or the central bank, we have from (3):

\[
\begin{align*}
  r^* &= \hat{G}(i; K, Y, z) := G(i) \\
  \text{(4)}
\end{align*}
\]

It is not unreasonable to assume that the mapping \( G(i) \) is increasing, an assumption we make below.

The significance of the required rate of return becomes evident when we combine the asset market equilibrium with the real economy (producing firms). The question then becomes: will an active monetary policy, which determines the required rate of return \( r^* \), be compatible with macroeconomic stability; i.e., a stable flow of private investments, smooth private consumption, and a high and stable employment level? As will be demonstrated, Haavelmo’s answer is “No!”. Asset market equilibrium requires real capital to earn a return equal to the required rate of return, \( r^* \). At the same time, capital equipment is rented by producing firms at a price \( r_K \) and used to produce some aggregate output according to the macro production function \( \phi(K, N) \), with \( N \) as aggregate employment.\(^\text{12}\) With \( w \) denoting the real wage rate, and with a fixed price level, real profits are \( \phi(K, N) - (\delta + r_K)K - wN \), where \( \delta \) is a fixed rate of depreciation. In equilibrium, for fixed \( K \), we must have \( r_K = \phi'^*_K(K, N) - \delta \). The real rate of return, \( \phi'^*_K(K, N) - \delta \), or Wicksell’s productivity rate as it was dubbed by Frisch (1952), must determine the rental price \( r_K \). However, as long as the interest rate \( i \) is chosen by the monetary authorities, there is no reason to believe that this productivity rate of return will coincide with the required rate of return \( r^* \). For this reason capital owners will not the required rate of return from renting out their capital equipment, we can have \( r_K > r^* \) and \( r_K < r^* \). Hence if we demand economic explanations to be based on a determined mathematical model, we have a problem: the model we have formulated cannot explain how the economy functions in a regime with active monetary policy. The model is fundamentally overdetermined if \( r_K = r^* \) is imposed as an equilibrium condition.\(^\text{12}\)

\(^{12}\)This function has standard neoclassical properties: constant returns to scale, strictly increasing and with diminishing returns to each factor of production – hence the inputs must be technical complements.
3 Modified investment behaviour

Does overdeterminacy imply that active monetary policy is undesirable? Not necessarily. Overdeterminacy is only a feature of an economic model, not of the real world. Hence, the only mistake we can make is to use a wrong – or irrelevant – model to aid monetary policy. Specifically, models that implicitly or explicitly assume that capital markets are in joint equilibrium have low relevance for monetary policy according to Haavelmo:

It is obvious what an actual economy does under such circumstances: it operates under a different model that has a solution. Why, then, should we take even the slightest interest in an overdetermined model? If we do, the only acceptable reason would seem to be that we believe that, somehow, the economy first "tries out" the hopeless model, and then derives a practicable alternative in a way which could be predicted by studying the overdetermined model.

[Haavelmo, 1960, pp.200–201.]

In a discussion that followed after this passage, Haavelmo pointed out that one possible solution to the overdeterminacy problem could be to add a Wicksellian cumulative process to the model, a route he pursued in his inflation theories. However, for his business cycle model, the modification he proposed was to exclude $r_K = r^*$ as an equilibrium requirement. To close the model, he instead sketched a theory of investment behaviour when this equality fails to hold. This was a natural way for Haavelmo to attack the issue, since a state of disequilibrium plays an important role in explaining investment behaviour in his 1960-treatise. A point he stressed was that in a neo-classical model, there is no way to derive the demand for investment from the first-order condition with respect to capital. If the rate of investment is the rate of change of the capital stock, it cannot be determined by the rate of interest itself, but by the rate of change in the rate of interest, see (Arrow and Kurz, 1970, p.74) who attribute this insight to Haavelmo (1960). The theorist therefore has to look for other reasons than a pure profit motive to get a formal theory of investment. For instance, one could introduce supply side constraints or time-lags in the production of capital goods. The way Haavelmo gets around the overdeterminacy problem in SMT can be viewed as a short-cut to the more complicated job of modelling supply side constraints.

Let us focus on the source of the “problem”, namely that the exogenously chosen interest rate interferes with equilibrium in the capital market. When the required rate of return corresponds to the actual rate of return ($r_K = r^*$), the households (which are the investors) have an implicit demand for investments, passively investing their savings. When $r^*$ fails to match the marginal productivity of capital, other investment responses become relevant. One theory suggested by Haavelmo is:

---

13 This is discussed in Anundsen et al. (2012), and we return to this issue in Section 5.
14 See Boianovsky (2002) for comments on how this influenced, and was influenced by, the unemployment theories of Don Patinkin.
15 Several authors have recognised these insights, see e.g. Nickell (1978, p. 12).
Investment demand = \begin{cases} \infty & \text{if } r_K = \phi'_K(K, N) - \delta > r^* \\ \text{Passive} & \text{if } r_K = \phi'_K(K, N) - \delta = r^* \\ -\infty & \text{if } r_K = \phi'_K(K, N) - \delta < r^* \end{cases} \quad (5)

This is a formal way of saying that when the money market interest rate is set too low, investors will try to purchase as much capital equipment as possible. When it is too high, they will not invest at all. The implication for investment demand is an infinite positive or negative rate when $r_K \neq r^*$, since investors want to reach the new optimal level instantaneously.\(^\text{16}\)

Infinite demand cannot be an equilibrium outcome. Since there is no capital price in this one-sector economy, there is no way for investors to signal their desire for more capital, and no reason for consumers to cut back on consumption to make a larger piece of the pie available for investments. We cannot use a theory of optimising behaviour to determine how total production will be allocated between consumption and investments. Haavelmo’s solution was to impose a rationing scheme for investors. This implies that the model can have a solution with excess demand, and he assumed that consumption has priority to the gross domestic product, i.e. the amount of private sector investments is residually determined. This will put an upper bound on the maximum supply of new capital goods equal to full employment GDP less consumption. The lower bound will be zero if we assume that capital cannot be transformed instantaneously to consumption goods. Realised investments are therefore:

\[
\text{Realised investments} = \begin{cases} \text{Rationed} & \text{if } r_K = \phi'_K(K, N) - \delta > r^* \\ \text{Passive} & \text{if } r_K = \phi'_K(K, N) - \delta = r^* \\ 0 & \text{if } r_K = \phi'_K(K, N) - \delta < r^* \end{cases} \quad (6)
\]

Haavelmo combined these different theoretical elements to formulate a short-run macroeconomic model which generates endogenous cycles, or maybe more to the point: secular changes with switches between periods of full employment and recessions.

4 A Business Cycle Model

We now outline the business cycle model of Haavelmo, where the implications of the fundamental overdeterminacy and the possible investment responses are taken into a dynamic macroeconomic model. The original presentation is found in Part VI of SMT. It is a Keynesian type macro model for a closed economy, where the investment response of firms in the economy plays a fundamental role. Haavelmo included a brief presentation of this model in an article on business cycles in the International Encyclopedia of the Social Sciences, see Haavelmo (1968), but the only full-fledged presentation is found in SMT. In this section we give a self-contained but condensed presentation.\(^\text{17}\)

\(^{16}\)This point was stressed on several occasions in Haavelmo (1960).

\(^{17}\)A more detailed presentation of the model can be found in Anundsen et al. (2012).
Haavelmo’s theory integrates the market for real balances and the hypothesised investment behaviour outlined above, with standard building blocks of i) an aggregate production function with labour and capital as inputs, ii) a Keynesian consumption function, and iii) an assumption of a given labour supply.

It is convenient to assume that the law of indifference relationship (4) gives the required rate, \( r^* \), as an increasing (and differentiable) function of the money market interest rate, \( i \):

\[
r^* = G(i), \quad G'(i) > 0
\]  

(7)

First, let us consider the momentary equilibrium. There are two relevant situations to consider; one with \( r_K < G(i) \) and one with \( r_K > G(i) \). As will become obvious from the discussion of the dynamic equilibrium, a situation with \( r_K = G(i) \) cannot be maintained as stable equilibrium. The reason is that the dynamics of the two inputs will cause the law of indifference to break down at some point in time.

A situation with \( r_K < G(i) \) goes together with zero investments. The households do not achieve their required rate of return, and would – if they could – get rid of capital equipment. Of course, gross investments cannot be negative, but this will result in a negative net investment due to depreciation of the existing stock. This is in accordance with the discussion in Section 3. Aggregate demand is then determined by private consumption as well as government expenditures (for later reference denoted \( E_G \)). We refer to this situation as a “depression”, and it involves only a minimum level of employment, \( N_{min} \), resulting in unemployment.

The other case with any persistence, is \( r_K > G(i) \). Since the household sector earns a higher return from holding real capital than they require, they will invest as much as possible in order to increase their stock of capital equipment. In this case, aggregate demand exceeds the production capacity, and either consumers or firms have to be rationed. As noted above, consumption gets priority. This implies that private investments are given by full capacity GDP less private consumption and government demand. It follows that in a period with \( r_K > G(i) \), there will be full employment (\( N = H \)) and no idle productive equipment. We refer to this as a “boom”.

Labour supply is modelled by assuming that the entire labour force is willing to work as long the wage they receive exceeds some reservation level.\(^{18}\) During periods of unemployment, competition on the supply side drives the wage down to the reservation level. Under full employment, competition on the demand side will, on the other hand, push the wage up to the marginal productivity of labour.\(^{19}\)

\(^{18}\)The reservation wage will fall below any marginal productivity of labour at full employment.

\(^{19}\)Note that this requires us to think of the firm as changing strategies, from being a price taking profit maximiser in booms, to produce to order during recessions.
Figure 1: The left panel of the graph shows the actual rate of return to capital as an increasing function of employment (consistent with inputs being technical complements). The right panel shows the required rate of return as an increasing function of the nominal interest rate. For an initial level of employment $N_a$, we have $r^* = G(i_0) < r_K$.

Figure 1 illustrates an initial situation with unemployment ($N_a < H$) and a given capital stock. For a fixed money market interest rate, $i_0$, the required rate of return follows, whereas the initial level of employment pints down the actual rate of return (equal to the net marginal productivity of capital). Clearly, given the premises above, the situation illustrated in Figure 1 cannot constitute an equilibrium. It will immediately be replaced by a situation with full capacity utilisation and full employment, since as much as possible will be invested when $r_K >^*$. As shown in Figure 2, this will – since labour and capital are technical complements – increase the actual return for a given level of the capital stock. This situation will persist, and will stimulate new investments. However, as time passes, the increase in capital stock will make the curve in the left panel shift downwards. Hence, if the interest rate is fixed at $i_0$, we will at some point in time get to a state where $r^* > r_K$. 
Figure 2: Momentary equilibrium in the case where \( r^* = G(i_0) < r_K \)

Figure 3 represents an initial situation where the required rate of return is higher than the actual rate, and the theory predicts zero investment activity and unemployment. Hence an employment level such as \( N_b \) cannot be an equilibrium. Instead we end up at \( N_{min} \), in a regime with economic depression. By the same logic as above, this situation will be quite persistent, since the reduction in employment makes the rate of return drop even further below the required rate. In addition, it is likely that this situation is even more persistent than the one with \( r_K > r^* \). The reason is that the only way the economy can work its way out of unemployment, is through an increase in the actual return, which is accomplished by a gradual depreciation of \( K \). This process can take a long time. As the capital stock is reduced, the curve in the left panel is gradually shifted upwards. Eventually, we get \( r^* < r_K \), for which the relevant illustration is again Figure 2.

Figure 3: Momentary equilibrium in the case where \( r^* = G(i_0) > r_K \)

We label the two regimes A and B, respectively. The theoretical framework then implies that the economy will be unstable and that it switches between periods with
(mass) unemployment and periods with full capacity utilization.

Assume for simplicity that both the money market interest rate and the full employment level are constant. Under appropriate assumptions for values of the various exogenous variables, we have that:

- Under Regime A, the marginal productivity of capital *declines* because of positive net investments
- Under Regime B, the marginal productivity of capital *rises* since the capital stock is gradually depreciated

Suppose that that we start out in regime A at some point in time, $t_0$. With full capacity utilisation, firms are investing as much as they can. However, as capital is accumulated, the marginal productivity of capital will decrease – this is illustrated in panel (a) of Figure 4. At $t_1$, it has fallen to a level equal to $r^*$. As soon as the marginal productivity falls below this rate, there will be a switch to regime B. Firms stop investing, leading to a sudden drop in output and a jump in the unemployment rate. Further, since the stock of capital is unchanged even though employment has fallen, the marginal productivity of capital jumps down to a level far below $r^*$. These shifts are depicted in the three panels of Figure 4, which for authenticity is the same illustration used by Haavelmo (1966, p. 158). After the sudden drop, the marginal productivity of capital recovers as the stock of capital is worn out. When we reach $t_2$, the actual return has again become equal to $r^*$, and as soon as it is marginally above $r^*$ we switch back to regime A. Firms start investing again, leading to a jump in production and employment. As a result, the marginal productivity of capital jumps up, and a new cycle is initiated. Due to coordination failures, the investors cannot fine-tune their purchases of new capital, and this prevents $r_K = G(i_0)$ from being stable.20

---

20It is interesting to note that more than a decade later, Malinvaud (1980) suggested similar ideas for macroeconomic modelling. In a footnote on p. 10, in his 1980-book, Malinvaud wrote “The model to be developed here has a formal characteristic that can also be found in some of those elaborated by the theory of business fluctuations, namely, it contains several “regimes” and it views the economic evolution as shifting at times from one regime to another. But each one of its regimes will be based on a complete economic analysis and will claim to represent, although in a very simplified manner, a type of situation that can be observed for a more or less prolonged period.”
Figure 4: Haavelmo’s Business Cycle Model. See text for explanation.
At this point it is worth remembering that the law of indifference in the capital market \((r^* = G(i))\), is generally a multivariate relationship, represented by the variables collected in the vector \(z\) in (4). For instance, it is reasonable to assume that the required rate depends on transaction costs and liquidity constraints (see Chand (2012) for an application of this kind). In a situation with low liquidity, with little trust and with lack of confidence in the financial system, the function \(G(i)\) is shifted downwards for a given level of the money market interest rate, \(i\). Experience suggests that there is no guarantee that a reduction of \(i\) gives the same reduction in the required rate as before liquidity dried up. A financial crisis can lead to a full fledged depression within this framework.

More generally, what matters is not really the exact time-profile of the marginal productivity of capital, but how this time profile matches the path of \(r^*\). In SMT, it is shown that both A and B are highly persistent, and that the case of symmetric cycles illustrated above is a special case, meaning that Figure 4 is a crude simplification regarding what dynamic pattern the model can produce. Haavelmo stated that:

The importance of the interest rate for continued economic expansion is not determined by whether it is high or low, whether it is adjusted upwards or downwards etc., the whole point is to keep the inequality in favour of the marginal productivity of capital.

[Haavelmo, 1966, p.153, our translation.]

In line with this wider interpretation, Haavelmo’s theory is more like a model of secular cycles, or intermittent depressions of irregular and unequal lengths, than an ordinary business cycle model. The point at which the economy changes from a boom to a recession, he labeled the “point of catastrophe”. With grim realism he stated that, although we enjoy full employment and stability a catastrophe may be just around the corner.

5 Relation to theories of his time

The regime-switching process and the distinct boom-recession-boom cycle of Haavelmo’s model results from his two main premises. First, he replaced the notion of a mathematically well-behaved (smooth) investment function with the idea of arbitrage-based investment strategies. Second, he showed that in a model that included real capital as one asset investors can hold, the money market interest rate cannot be autonomously controlled by the central bank without causing out-of-equilibrium behaviour. Haavelmo’s logical scheme implies that imbalances between supply and demand for capital and other assets may occur. When applied jointly in a macroeconomic model, the two principles imply disequilibrium macrodynamics with endogenous switching between low activity and full employment regimes. Taken at face value, Haavelmo rejected the Neo Classical Synthesis macro model, which was the standard approach up to the stagflation period that followed in the wake of the two OPEC oil price shocks in the 1970’s.

In formulating a full fledged macroeconomic model with endogenous switching between a regime with classical and smooth growth and another with Keynesian unemployment, Haavelmo preceded the disequilibrium, or fixed-price, macroeconomic models of the 1970s and 1980s; see e.g., Barro and Grossman (1976), Malinvaud (1977), Bénassy
It is interesting to observe that while these models first abstracted from investments and capital markets, the core idea of Haavelmo’s model is that disequilibrium constellations arise in the interface between asset markets and the real economy. The profession seems to a large extent to have lost interest in disequilibrium macroeconomic models around 1990. It will never be known whether or not the perspective adopted by Haavelmo in SMT could have provided a powerful guideline and a different direction of development.

Moene and Rødseth (1991) pointed out an interesting parallel to Tobin’s q-theory of investment, see Tobin (1969). The difference is that while Tobin’s theory needed rationalisation in the form of convex adjustment costs to obtain “smooth” investments, as in Hayashi (1982), Haavelmo – in an extended two-sector version – obtained a logically sound solution for the investment level with reference to the productive capacity of the macro economy itself. Hence, because of the way investments are determined, Haavelmo’s theory can be said to be supply-side oriented in both the two-sector version, and in the “cruder” form presented above.

The preceding sections have demonstrated how Haavelmo’s macroeconomic business cycle model has Keynesian features, and that it was greatly inspired by the work of Knut Wicksell. Both earlier work, and work published around the time Haavelmo developed these ideas, such as Goodwin (1951), Stein (1969, 1970) and Fischer (1972), do – in some respects – bear a resemblance to his business cycle model. The rest of this section is devoted to a discussion of both the heritage from Keynes and Wicksell, as well as the similarities and differences between the model of Haavelmo and other theoretical approaches suggested around the same time.

### 5.1 Heritage from Keynes and Wicksell

As mentioned above, Haavelmo was an adherent of Keynesian theory and Wicksellian analysis. His interest in the work of Wicksell was influenced by Frisch, who gave lectures on Wicksell in the 1930s, when Haavelmo was a student.

In Haavelmo’s model, the standard effects of Keynesian liquidity constraints are replaced by the effects of having the rate of return being different from the required rate, i.e., \( r_K \neq r^* \). Recessions are therefore caused by periods in which the real rate of return falls below the required rate. This is different from the Keynesian mechanism, where changes in liquidity preference and lack of effective demand lead to a rush to liquidity, causing a drop in investment and output. Instead, because of the role played by the law of indifference, it can be more relevant to regard Haavelmo’s model as an extension of those pre-Keynesian theories of the cycle that were bearers of a Wicksellian influence.

As shown by Boianovsky and Trautwein (2006), one important example is the work on business-cycle theory by Swedish economist Erik Lindahl, where the difference between the market rate of interest and the normal interest rate is a central concept. There are, however, important differences, for example Lindahl’s theory, as we understand it,

---

21Some thoughts on how to analyse behaviour outside equilibrium, were formulated by Haavelmo (in Norwegian), in 1958, titled “What can static equilibrium models tell us?” (Haavelmo, 1958a). Arrow published similar ideas at the same time, see Arrow (1959). Axel Leijonhufvud, translated Haavelmo’s 1958 offering, and was instrumental in getting it published in Economic Inquiry in 1974.

relies primarily on a super/subnormal money market rate of interest, to generate cycles.

Our interpretation of Haavelmo on this point is supported by his own discussion in SMT, where he argues that his business cycle model can be used to provide an alternative explanation of the Keynesian liquidity trap (the situation where changes in monetary policy have no effects). In Haavelmo’s model, a “liquidity trap” occurs because investors’ required rate of return exceeds the actual return, no matter what the policy rate is. Such a case is illustrated in Figure 5, also taken from SMT. For the required rate function depicted, any change in the policy rate will be ineffective as long as the economy is in a state of recession, i.e., with $N = N$. The policy that may boost economic activity is fiscal stimulus, which raises employment to $N = N'$. Thus, the return to capital in the low employment regime will increase as well. Such a policy is illustrated by the shift initiated by some increase in government expenditures, $\Delta E_G$ in Figure 5.

![Figure 5: Potential liquidity trap](image)

It is possible that this fiscal stimulus may shift the economy back to the boom regime for some low level of the interest rate.

In our presentation of Haavelmo’s boom-recession model, we have been faithful to the original formulation which abstracts from price dynamics and expectations of price changes. At one level, the assumption about a fixed price level is just a convenient simplification, and it may reflect the view that even if it was included, price level dynamics would not eliminate the boom-recession cycle. In the depression regime, the assumption may be quite realistic, because of downward nominal rigidities. However, for the full employment regime, with $r_K - r^* > 0$, it becomes more dubious to abstract from price increases even if capital accumulation, as noted, will also increase production capacity and alleviate some price pressure through that channel. Therefore, in this regime, it lies close at hand to specify a cumulative process of the Wicksellian type:

$$\pi = \theta (r_K - r^*), \theta > 0$$

with $\pi$ as the inflation rate, and $\theta$ as a parameter that measures the speed of the cumulative price process. If $\theta$ takes a low value, because of a capacity increase, international trade, or political and institutional innovations, the cumulative process may become negligible. Based on post WW-II experience, with relatively stable prices compared to other
post-war periods, Haavelmo may have thought it tenable to abstract from the cumulative process in his business-cycle model. Moreover, there is nothing in his model that makes \( r_K = r^* \), even with an added inflation process. Finally, in a situation with unemployment, Haavelmo would probably be reluctant to use (8) as a model of price dynamics. To use a central concept from the “Probability approach”, this relationship is not autonomous with respect to the employment situation in the economy. In fact Haavelmo had noted that either Wicksellian or Keynesian responses might be the consequence of departures from the law of indifference in the capital market:

It only remains to express what the reader must already have observed, viz., the extremely close connection that exists between the “Keynesian” effects of liquidity constraints and the Wicksellian theory of the cumulative process [in the model discussed]. The main difference is perhaps that one feels a little more comfortable about the Wicksellian theory when \( \rho^* < \bar{\rho} \) and a little more comfortable about the Keynesian ideas when \( \rho^* > \bar{\rho} \). It is hard to tell what is the proper verdict on the case \( \rho^* > \bar{\rho} \) because this case may actually be far more unstable with regard to producers’ behaviour than it appears from our scheme above

[Haavelmo, 1960, p.208-9.]

These observations – how Haavelmo viewed his “switching mechanism” as a substitute for liquidity constraints and his ideas for price dynamics – clearly illustrate the Keynesian and Wicksellian influence on his macroeconomic thinking. Interestingly, as documented by Boianovsky and Trautwein (2001), Wicksell seems to have closed the gap between his cumulative process theory and his “real business cycle” theory. Boianovsky and Trautwein show that Wicksell’s lecture notes from 1902/05 suggest that crises can result from differences between the expected real interest rate (corresponding to \( r_K \)) and the market interest rate. As we have seen, Haavelmo built on the same insight, but he may have downplayed the force and relevance of the cumulative price process for post WW-II Western economies too much, as the “creeping” rise of the price level changed nature into persistent inflation in the course of the 1970s and 1980s.

The investment demand in (5), may perhaps be seen as a quantity-version of the mechanism underlying the Wicksellian Cumulative Process. One version of this process is outlined in chapter 33 of Haavelmo (1960). A positive gap, in particular, between the productivity rate and the required rate of return will, in the short-run, trigger a movement in absolute prices. In the medium-run, it will result in a change in quantities. One question is how the changes in prices might affect the rate of desired capital accumulation. In the business cycle model of Haavelmo, he rules out price changes; not because he found price changes unimportant, but he wanted to focus solely on the dynamics of real variables, and perhaps he found his quantity-mechanism more relevant in the medium-run perspective. As noted by Haavelmo (1960; p. 208), in the case with a desire for “infinite investment”, capital accumulation might itself tend ”to slow down the rise in prices”. We might conjecture that even if we should have introduced price dynamics, the underlying impact on the desire for capital accumulation will not be changed fundamentally, perhaps, only modified.

\(^{23}\)In this quotation, \( \rho^* \) corresponds \( r^* \) and \( \bar{\rho} \) corresponds to \( r_K \) in our notation.
5.2 Goodwin and Haavelmo

Like many business cycle models developed in the 1940’s and 1950’s, Haavelmo’s model relies on assuming that the investment function is non-linear. Another contemporary business cycle model was the one developed by Goodwin, see Goodwin (1951). This theory suggests that net investments are zero if the actual capital stock is equal to the “desired stock of capital”, which again depends on the activity level in the economy. Net investments are positive if the actual capital stock is less than this desired stock. They are negative if the actual stock is greater than the desired stock (due to the depreciation of existing capital equipment). The similarities between the two theories are apparent; both are self-contained and do not rely on outside disturbances to generate business fluctuations. Cycles are generated by endogenously determined shifts in investment strategies.

In contrast to Goodwin, who assumed that there exists an upper limit on the desired capital, Haavelmo made an attempt to give behavioural content to the theory of investment demand. In SMT, it is clear that he was critical to Goodwin’s approach, which he argued disconnects the desire for more capital and the profitability of new investments. A short way to put it is that it ignores the relation between the marginal return to capital and the required return. As we have seen, incorporating the financial markets is crucial to the dynamics of the business cycle model suggested by Haavelmo. A similar point has been stressed by Zarnowitz (1985), who criticises the Kalecki (1935), Kaldor (1940), Goodwin (1951) type of models for ignoring the monetary and financial side of the economy, “[...] which theory and evidence suggest, are particularly important in major cycles, crisis, and depressions” (Zarnowitz, 1985, p.541).

In SMT, Haavelmo’s own critique of the Goodwin type of models is formulated rather explicitly:

> At one level of analysis, the theory developed by Goodwin and others seems rather convincing. It provides a mechanism that is able to explain the puzzle of business cycles as a necessary consequence of the development of the economic system. These mechanisms are, however, problematic to reconcile with real-life behaviour. In the theory of Goodwin, the entire problem is artificially reduced, or eliminated, by assuming certain limits to the producers’ desire for capital, which is not well founded in behavioural theory.

[Haavelmo, 1966, p. 156, our translation.]

It is clear from this quotation that even though Haavelmo acknowledged the Goodwin model to be self-contained, he also noted its limitations in terms of missing behavioural content. By taking into consideration the alternative uses of capital or wealth management, Haavelmo thought he had closed this gap, since business cycles in his model are driven by investors’ profitability considerations and not *ad hoc* upper and lower limits on the demand for real capital.

5.3 Linkage to modern Keynes-Wicksell models

Another interesting comparison is the one between Haavelmo and the Keynes-Wicksell (KW) monetary growth models that were developed in the 1960s. References to this literature are Stein (1969, 1970) and Fischer (1972), where the prototype KW model
is compared with a neoclassical version. Their view is based on the fact that in KW models, planned investments and savings are determined by different agents and therefore represented by different functions. In contrast, neoclassical models have savings defined as planned investments.

In Haavelmo’s model, the main determinant of the level of economic activity is the difference between the actual rate of return, $r_K$, and the required rate, $r^*$. Translated into levels of capital, it can also be understood as a relationship between the ‘optimal’ stock of capital (for some level of employment), $K^*$, and the actual stock, $K$, i.e. giving behavioural content to Goodwin’s “desired stock of capital”. The abrupt shifts between booms and recessions in this model are caused by the desired rate of investment being ill-defined, since investors ideally want to reach $K^*$ immediately. If this is to happen, an infinite rate of investment is required. If, on the other hand, the optimal stock of capital should fall below the existing stock, desired gross investments could be negative, but Haavelmo only allowed net investments to be negative in his model. The investment function implemented in the model combines the desired investment level of investors with the upper and lower bounds:

\[
\text{Investment} = \begin{cases} \text{Rationed} & \text{if } K^* \geq K \\ 0 & \text{if otherwise} \end{cases}
\]

(9)

In KW models, on the other hand, it is assumed that the realised level of investment is a weighted average of household’s planned savings and investors’ desired savings, no matter what the relationship between $K^*$ and $K$ is.\(^{24}\) The level of desired savings is a linear function of $K^* - K$, while planned savings are income net of consumption. In light of Haavelmo’s main point, this is problematic since it does not take into account the law of indifference. Capital keeps accumulating, even in periods when investors would have preferred less capital.\(^{25}\)

The symptoms of disequilibria are therefore quite different. In Haavelmo’s model investment will switch between two regimes. This is a mechanism missing from KW models. Instead these models include a Wicksellian cumulative process that produces price changes proportional to $K^* - K$. Indeed, this process is similar to what Haavelmo proposed in his 1960 book, providing further evidence of the similarity between their approaches.

Hence, KW models and the business cycle model of Haavelmo are hard to distinguish under the full employment regime (when $r^* < r_K$, or $K^* > K$). Relying on the Haavelmo-story, we can tell a more convincing story, by introducing a depression regime for the case when investments logically fall to zero when $K^* < K$. Regarding this difference, one should bear in mind that the KW models were designed with a different objective than Haavelmo’s. Indeed, Stein states in his 1969 article that “[r]ealistically, we should relax the assumption of full employment when there is price deflation. But for reasons of simplicity [...]”. It is clear that they were concerned with monetary growth models. Nevertheless, Haavelmo’s considerations could possibly have been an important contribution to the Keynes-Wicksell literature.

\(^{24}\)The weights are supposed to allow for a general rationing scheme.

\(^{25}\)By modifying the investment function into Haavelmo’s non-linear/piecewise linear types, the two models could have been merged, making KW models consistent with the law of indifference.
6 Final discussion

In this paper, we have presented Haavelmo’s monetary theory of investment and business cycles, and we have presented its relationship to contemporary theory and the linkage to Keynes and Wicksell. A central issue in understanding how a market economy with decentralised portfolio decisions will operate is the problem of overdeterminacy related to the law of indifference in the capital markets. This relationship has interesting implications for the analysis of active monetary policy through interest rate setting. Even though issues related to the law of indifference were published already in Haavelmo (1960), the business cycle implications of this law were not spelled out in full until SMT. In this treatise, he presented an original, macrodynamic disequilibrium model with endogenous business cycles, where the activity level switches between a constrained full capacity-regime and a constrained low demand-regime.

As we have seen, the two assumptions of liberalised financial markets and a politically determined money market interest rate were important for the cyclical solution of his model. When the deregulation of Norwegian credit markets was completed in the 1980s, Haavelmo was sceptical to the dominant view that the interest rate could be used as an instrument for activity regulation. Interestingly, he referred to his macro model when lecturing “On the role of monetary policy in a deregulated credit market” in Norges Bank in 1987 (Haavelmo, 1987). This was in the middle of a difficult period for the Norwegian fiscal and monetary authorities. Haavelmo’s main message in the lecture was that the monetary policy regime, where the interest rate was used as an instrument to keep the exchange rate fixed, added to the problems by creating imbalances in the capital markets. Haavelmo did not get the response he had hoped for. When the Norwegian banking sector later collapsed, see for example (Reinhardt and Rogoff, 2009, p. 377) and the unemployment rate rose to a level unheard of since the 1930s, he commented with noticeable regret: “It turned out just like the theory predicted”.26 This example shows that although he was always a scholar, he saw his theoretical framework as relevant for practical policy thinking. Not at the level of daily operation though, but definitively as a guideline for choosing the optimal monetary and fiscal policy. Bårdsen and Nymoen (2001) argue that Haavelmo’s concern about unintended consequences of using the interest rate as a policy instrument is still relevant for inflation targeting central banks of our day.

SMT, which has been our main reference, was the result of his efforts during a sabatical in 1966. Why did Haavelmo only publish it as a textbook in Norwegian? Should this indicate that he did not have high thoughts about it? There is however no direct evidence saying that Haavelmo was unhappy with the product.27 As we have just seen, he referred to his model in discussions of the premises and consequences of the deregulation of Norwegian credit markets, that he gave on several occasions in the 1980s. It would be out of character if he at the same time thought poorly about his theory.

If one should speculate more about the underlying reason why Haavelmo did not publish in English, one might point to his personality, and the way his professional life evolved. After the ‘Probability Approach’ and throughout the 1950’s, Haavelmo was a...
star of the international economics profession. He was active in the international network of economists, was elected president of the Econometric Society, and he received an offer to move to the University of Chicago in 1955. In Norway, Haavelmo was held in high esteem among the top politicians and bureaucrats who laid down the path of economic and political development. This recognition must have been gratifying to him.

With the “Probability Approach”, Haavelmo had shown once and for all that he could raise a new and important research question and find a very satisfactory answer where no answer previously existed, see Bjerkholt (2007b). In “A Study in the Theory of Investment” (Haavelmo, 1960), he shoulder a task of the same intellectual proportions, but the reaction was more lukewarm. When the time came for his macro theoretical project in 1966, he may have trusted his own judgement that it was a reasonable good offering, but that the extra effort of touring with his results was perhaps not worth it. He may have felt that the macroeconomic chapters in the investment theory book, though short and terse in style, had made the core ideas available to the international profession. Besides, there was simply too much fishing to be done. As he expressed it in an interview with a main Norwegian newspaper in 1989: “I have had fun as an economist, but now it is far more important for me to go trout fishing.”28

28 Aftenposten, October 12 1989, p.17, our translation.
References


