

Trond Andresen

On the Dynamics of Money Circulation, Creation and Debt – a Control Systems Approach

Thesis for the degree of Doctor Philosophiae

Trondheim, October 2018

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics



Norwegian University of
Science and Technology

NTNU

Norwegian University of Science and Technology

Thesis for the degree of Doctor Philosophiae

Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

© Trond Andresen

ISBN 978-82-326-3398-2 (printed version)

ISBN 978-82-326-3399-9 (electronic version)

ISSN 1503-8181

ITK-report number: 2018-15-W

Doctoral theses at NTNU, 2018:304



Printed by Skipnes Kommunikasjon as

Some errors in original corrected, See appendix C for corrections done.
Any response appreciated. E-mail: trond.andresen@ntnu.no

Summary

This thesis develops and explores a way of understanding a macroeconomy; as a set of "nodes" interconnected via a network of money flows. A metaphor for this is a body with a circulatory blood system, furnishing the different organs with the necessary oxygen and nourishment. Later on in the thesis we will study the effects of indebtedness, which – with the "body" metaphor – may be compared to increasingly draining blood from the organism and sending it back in again - but as a non-flexible, damaging process.

This thesis does not concern itself with issues like the balance between consumption and investment, or what type of investment, or with the foreign sector. The view is that if a nation's economy (the "body") is allowed sufficient and flexible flows of money ("blood") to all its parts, it will "make do", and then one may choose somewhat different policies on top of that without ending up in serious crises.

A prerequisite for the analysis to follow is that the country in question issues its own currency (it is not in the eurozone or dollarised).

The main (hopefully useful) contributions of this thesis are as follows:

1. An understanding of time lags and time dispersion of flows in monetary circuits, and how this relates to money velocity and liquidity preference.
2. An aggregation approach that gives an alternative and straightforward connection between micro- and macroeconomic dynamics. This could perhaps be useful in the persistent discourse within the economics field about "microfoundations for macro".
3. A separate analysis of the advantages of time-continuous models, as opposed to the commonly used time-discrete models in stock-flow economics. All models are stock-flow consistent and achieve correct accounting – which is easier to ensure in continuous time.

4. It is argued that debt and money growth rates from commercial banks following a Bank of International Settlements (BIS)-type capital adequacy restriction will be inversely proportional to the mandated minimum capital/asset ratio. This is shown mathematically, and is a further and more precise contribution to the increasing (verbally expressed) recognition in the economics field that commercial (licensed) banks create money when they lend. It is shown that if banks only lent out of their net interest income, the money growth rate would be a magnitude lower than what empirical data say – while the "inversely proportional" rule gives a growth rate of a magnitude that corresponds well with empirical data. It is demonstrated how a theoretical new bank can bootstrap itself from a start only with the new owners investing their High Powered Money ("HPM"), and then continue with issuing an exponentially increasing loan flow, accompanied by the creation of credit money *ex nihilo* at the same rate.
5. It is argued that credit money growth via licensed banks is a *good* thing – in the sense that a system where all debt growth were accompanied by net money growth at the same rate, would not end up in debt crisis. Non-bank financial institutions (NBFIs) which extend loans, *but do not create money*, are the problem.
6. A distinction between "non-discretionary" (forced) and "discretionary" (more or less self-decided) money outflows from agents is introduced, and exploited in a crisis model.
7. A new stock-flow model, and simulations of debt-related crisis collapse mechanisms, are introduced and explored. Financialisation is represented as a slowly time-growing parameter, inspired by Minsky's Financial Instability Hypothesis. An increased debt burden leads to more insolvencies, leading to banks and NBFIs holding back in giving new loans. This again reduces flows in the real economy, in the next round leading to pessimism and increased liquidity preference among firms, capitalists and households.
Increasing liquidity preference is represented by the time lags of these groups (subsystems) in the model. This is inspired by J. M. Keynes' and Irving Fisher's writings. New loans not being given and increased general liquidity preference result in a damaging positive feedback that brings the system into crisis and collapse.
8. In the crisis model, the government is "embedded" as nodes and extra

parallel flows in the circulatory real-economy network. This enables an inclusion of the government debt burden together with the private real sector debt burden. (Internal financial sector debt – which net to zero – is ignored in this thesis.)

9. The advantages of a purely *electronic monetary system* is argued. There is no longer a need for bank branches for anything but vetting of potential borrowers (and perhaps giving advice on what sort of savings to choose – among these a spectrum of paper that could be offered at the Central Bank). Depositors can all have all types of accounts at the central bank. (Or in a private bank, but then being just as a mirror of a corresponding account at the CB.)
10. Electronic money enables a system where all circulating money is Central Bank money (HPM, "base money"), and thus cannot be lost. Such a system is proposed and described. It also has the property that only bank owners take the hit when a bank becomes insolvent. The system does not need government and Central Bank bailouts, and is extremely robust against crisis. Depositors' money cannot be lost, hence there will be no bank runs in such a system.
11. Modern Mone(tar)y Theory ("MMT") is supported. The control tool proposals in this thesis depends completely on the country issuing its own currency, also central to MMT. Dollarised or eurozone countries cannot implement the proposals. But see point **13** below.
12. It is explained how electronic money for the first time enables precise control against inflation and deflation, in contrast to today's sole, blunt and slow-impact interest rate tool.
13. Electronic money can be introduced easily, fast and cheaply. Systems where the main transaction tool is the mobile phone, are up and running – not the least in some poor countries – with an excellent track record. One could easily introduce a parallel (also called "complementary") electronic national currency in a dollarised or eurozone country that suffers from a debt-related crisis, and this will give a way to solve problems. This is explained and discussed.

My own publications that are relevant to this thesis are found in the reference list and are also referred to in the text when appropriate.

Preface

This thesis is dedicated to colleagues, political leaders and activists world-wide working for a modern, rational and humane monetary system, for simplifying and paring down today's bloated and parasitical financial sectors, and for nations' right to decide their own economic policies and try new solutions.

I am especially thankful for four years of interaction with the the Correa government in Ecuador, where they were pioneers in implementing a system with everyone being offered a deposit account directly at the Central Bank, and transactions via the mobile phone system. They were truly inventive trying to circumvent the problems of being a dollarised country. Recently however (from 2017 onwards), traditional forces have sadly taken over and reversed many things, but I hope for a comeback for a progressive reform government there.

Finally, I dedicate this thesis to my wife Solveig, and children Olav, Eirik, Gro and Laila.

Trondheim, July 2018

Trond Andresen

Contents

Summary	iii
Preface	vii
1 Introduction	1
2 A Signals- and Systems-based toolbox	5
2.1 Phillips' pioneering stock/flow model	5
2.2 Our approach	8
2.3 Why time-continuous models?	10
2.4 Lags, delays, velocity and dispersion	13
2.4.1 Dispersion in time	15
2.4.2 Impulse response vs step response	17
2.4.3 Agent dynamics with no inflow	17
2.5 On saving, "hoarding" and liquidity preference	17
2.6 Time lags for interconnected units	20
2.6.1 Micro to macro?	21
2.6.2 Definitions and introductory remarks	22
2.6.3 The mean time lag for units in series and in parallel	25
2.6.4 Mean time lag for a system with feedback	29
2.6.5 Mean time lag from a state space representation	29
2.6.6 Mean time lag via the step response	30
2.7 Application: an economy with households and firms	32
2.7.1 The economy simplified to a single first-order block	37
2.7.2 Velocity and existence of money in the circuit	40
2.7.3 A time lag block with a small T is simply unity	42
2.7.4 The system with a multiplier	43
2.7.5 A Kalecki-inspired application	46
2.7.6 An application: critique of a "dynamic" IS-LM model	50

2.8	Aggregating agents into a sector	56
2.8.1	An aggregation theorem	57
2.8.2	Allowing for differing agent transfer functions	63
2.8.3	Example: Lag of a sector with consumption goods, intermediate and investment goods	64
2.9	Aggregation check through Monte Carlo simulations	65
2.9.1	Network of units with rational transfer functions	65
2.9.2	Networks of units with irrational transfer functions	72
2.9.3	Conclusion	73
3	Discrete versus continuous time	77
3.1	A time-discrete model	77
3.1.1	A time-continuous and better model	81
4	Lending and financial accumulation	89
4.1	The basic moneylending structure	90
4.2	The annuity-type debt service subsystem	94
4.3	When may debt “explode”?	95
4.3.1	Pyramid game ("Ponzi") dynamics	98
4.3.2	Accumulation with annuity-type debt service	99
4.3.3	Firms with no income during a start-up period	103
4.3.4	Can capitalists earn any profits when money is debt- created?	104
4.3.5	An intuitive, brief description	105
4.3.6	Real growth does not save the system	106
4.3.7	A measure of financial fragility	107
4.3.8	Summing up	108
5	A modern financial system and dynamics	111
5.1	A generic bank model without a Central Bank	112
5.1.1	A realistic and interesting result?	114
5.2	Including a Central Bank	119
5.2.1	The bank model with a Central Bank and taxation	119
5.2.2	A remark on government bonds	121
5.2.3	Money growth pari passu with debt – a good thing!	122
5.3	"Bootstrapping" a money-creating bank	123
5.3.1	The bank as a control system	125

6	Debt build-up and crisis	129
6.1	A "Minskyan" long-term crisis model	129
6.1.1	Including a government and Central Bank	134
6.2	Simulation experiments	135
6.2.1	All time lags and loss rates constant, only varying c_{FI}	135
6.2.2	Increasing c_{FI} , also T_{KF} and π	136
6.2.3	Growing c_{FI} , profit share, time lags & loss rates . . .	138
6.2.4	Holding all time lags except T_{KF} constant	144
6.2.5	Constant c_{FI} i.e. unchanging financialisation	146
6.2.6	An endogenous c_{FI} instead?	146
6.2.7	More on money velocity	149
6.2.8	A simple graphic portrayal of the model	150
6.3	Further about debt crisis models	150
6.3.1	A suggested cycles, bubbles and waves taxonomy . . .	152
7	Reform proposals	155
7.1	The two unsolved problems	155
7.1.1	Avoiding debt burden growth	156
7.1.2	How to get out of a debt-induced crisis?	156
7.2	Possibilities with an electronic money system	157
7.2.1	The Central Bank and electronic money	159
7.2.2	The "heterodoxy" is partly hostile to 100% money . .	160
7.2.3	Ensuring that only bank owners take the hit	163
7.2.4	A BIS-type capital adequacy constraint still feasible .	163
7.2.5	The CB as a lending and savings hub for society . . .	165
7.3	Inflation and deflation control	165
7.3.1	Electronic money and MMT	165
7.3.2	A problem – injection and drainage asymmetry	168
7.3.3	M and v control	169
7.3.4	Fiscal policy and capital controls	172
7.4	Parallel electronic currencies	172
7.4.1	The proposal	173
7.4.2	How does it work?	174
7.4.3	Discussion	175
7.4.4	Summing up: far better than the bleak alternatives . .	177
7.5	Conclusion on electronic money – and the rest	177
A	Block diagram representation	185
B	Biblical quotes against interest	189

C Corrected errors in the original submitted thesis	191
---	-----

Chapter 1

Introduction

This thesis sets out to contribute to a toolbox for macroeconomics with approaches from the control systems field. Such tools are then applied to the study of the dynamics of financial accumulation, mirrored by increased general indebtedness – and crisis developing because of this. Reform ideas for regulation are finally proposed and discussed.

This work is among other things motivated by the following factors:

1. *The lack of recognition or interest among the academic economics mainstream (the "neoclassicals") for debt/indebtedness as a central system state in macroeconomic dynamics.* It is fair to say that this factor has been generally ignored. This is probably the main explanation for why such a large share of the profession did not see the latest global financial crisis coming, well documented in (Bezemer; 2009). This general ignorance – or perhaps: choice of not seeing – is all the more serious since the neoclassical academic mainstream has immense influence on public discourse and policy choices.
2. *The non-use in "heterodox" ("anti-neoclassical") academic economics of convenient mathematical tools for dynamical systems.* The good thing however, is that debt build-up and its dangers are very much recognised by many of these economists, primarily those associated with the Post Keynesian school. Such economists however, are politically much less influential than the mainstream. There is a large literature among them on the topic of debt build-up and excessive growth of the financial sector leading to crisis, with perhaps the most well-known researcher being the late Hyman P. Minsky (Minsky; 1992). Most of the analysis – however – is of a verbal type, which makes it

lacking in some respects. Moreover, it does not help that a general skepticism towards mathematical tools (mostly a reaction to mis-use or over-use in the mainstream) is quite common and considered *comme il faut* among the heterodoxy.

Admittedly, there are a few exceptions to this, like Godley and Lavoie's discrete-time stock-flow consistent monetary macroeconomic models (Godley and Lavoie; 2007) and Steve Keen's continuous-time business-cycles-with-debt-accumulation models (Keen; 1995). We will return to these.

3. *The rich collection of tools and insights from control theory and engineering practice*, which has not been adopted by many economists. A striking characteristic of economics seen from a control systems position is the lack of interest in and understanding of the concept of stability, which is crucial for a control engineer (although the global financial crisis seems to have changed this slightly for the better). Generally, there is skepticism from the mainstream in economics towards control action ("regulation") as such, which for a national economy must be executed by the government or some of its agencies. This may be seen as a reaction to the failures of historical attempts at planned ("communist") economies, but one cannot because of this deny the need to stabilise – and in a sufficiently good manner run – a complex dynamic system with inherent instabilities, which a macroeconomy arguably is. A well-run macroeconomy could be compared to a process plant or a power plant: when disturbances arise, the goal and control action is to uphold the system running at constant and optimal settings.

4. Related to point 1 and 3: *The recognition of the merciless mechanics of accumulation due to compounding of interest/returns, also warned against in different religious and other ancient writings* (see appendix B for some quotations). This insight, and corresponding repeated admonitions all the way back to antiquity, is in modern times vigorously promoted by Christian and Muslim groups, and by some left-wing activist (or "green") groups. It seems that because such activity occurs mostly outside university circles – on "the fringe" – this important phenomenon is kept at an arm's length by academics: not only by the economics mainstream, but also by many within the academic heterodoxy (one could perhaps call this the "orthodox heterodoxy"). To this writer this seems to be an example of throwing the baby out with the bathwater. An academic exception to this is Michael Hudson's excel-

lent work on the role and dangers of compound interest since antiquity (Hudson; 2000).

5. *The importance for society of economics, and the level of animosity between different schools and subcultures in academic economics.* Macro(and micro-)economics is not an exact science. Academic activity being inexact is in many fields not a big problem: the discipline of literature studies (or other esthetics-oriented disciplines) is also inexact, but disagreements between different schools are mostly just spice in the dish, and have no important negative consequences for society. Macroeconomics however, is not only descriptive. It is also (meant as) a science for real-time monitoring and control of dynamic systems, and extremely important ones at that (national and global economies). Consequences of large errors in control strategies (both regulatory frameworks and running policy decisions) may thus be severe.

The political-economic choices made by governments, supranational economic organs and other powerful entities are not always popular. But they are often implemented in spite of this. Therefore there is a need for "scientific explanations" of why such economic policy choices "must" be made. For this, academic economists are called for. They become elevated to modern days' version of the clergy in older times: telling the populace why certain sacrifices must be made and then things will be good later on, or in the afterlife. In academic economics, this crucial ideological role in public discourse may ensure tenure, secure funding and well-paid careers for those who adhere to the "right" school of thought, while others are marginalised. Therefore more is at stake in the academic economics profession compared to other disciplines. This partly explains the strong animosity between different schools.

6. *New promising technologies* that may be applied to improve monetary systems. One example is paperless ("electronic") currency with exchanges mediated via the internet and mobile phone networks. For the first time in the history of economics some qualitatively better ways to run monetary systems have become feasible, enabling reforms that have earlier been proposed by creative economics thinkers and dissidents. An early pioneer for alternative monetary thinking – whose ideas have now become technically very feasible – is Silvio Gesell, with his *Schwundgeld* concept (Gesell; 1958), tried out through the succes-

ful parallel currency experiment in the town of Wörgl in Austria in 1932 (Lietaer; 2010), and the related "Stamp Scrip" proposal by Irving Fisher in the U.S., as a measure during the great depression in the thirties (Fisher; 1933b).

7. *A pro-enlightenment view*; that it should be – and is – possible for human civilisation to progress, and that science and technology is the central factor for this. In this however, understanding of the mechanism of *the self-fulfilling prophecy* is crucial: For better futures to be realised, they have to be aired and discussed in the public sphere, and only by this they will gradually seem, *and then actually become*, more realistic (Andresen; 2003).

8. *Having an engineering background* and working in an environment where colleagues are studying and improving technical and industrial systems, I am very much aware that a high living standard in a society basically depends on the technical sophistication of its production, its communications-, transport- health- and other infrastructure – and a correspondingly educated population that is able to exploit and further develop this. One may ask whether it is not paradoxical with such a background to choose to focus on a non-material entity like money. The answer is that a well-functioning monetary and financial system is a fundamental prerequisite for a high-tech society to succeed. It may be considered a part of society's transport and communications infrastructure. And when this financial infrastructure is well-designed and well-managed, it may – just like a highly automated factory – be run smoothly with far fewer employees than today. One goal of this thesis is to contribute towards a society where the financial sector is shrunk to a lean but much more well-functioning minimum, also liberating personnel for more meaningful work. With a better understanding of the dynamics of money and finance, and of today's technical possibilities, this becomes a feasible future.

This thesis is cross-disciplinary and stands with one leg in the monetary macroeconomics field, and the other in the system dynamics/control systems field. A side effect of speaking to both camps is that parts of the material here may be trivial (or difficult) for the economist reader, while other parts may be trivial (or difficult) for the control engineer. This can hardly be avoided.

Chapter 2

A Signals- and Systems-based toolbox

A macroeconomy evolves through time, and may be viewed as a collection of connected sub-entities or "nodes" interacting with each other. It is thus a *dynamic system*. Dynamics are mathematically and conceptually much more complicated than statics, including the “comparative statics” commonly used in economics (and critiqued in subsection 2.7.6). When doing dynamics, algebraic equations (corresponding to the intersecting graphs ["schedules"] widely used in economics) are replaced by differential (or in the discretised time case: difference) equations. These equations are difficult to work with in the sense that one can hardly – as one can in a static framework – find graphic or algebraic (when possible) solutions to them without computer-implemented solution software. It is also difficult to gain any qualitative insights about the interactions governing a dynamic system by inspecting the set of its differential (or difference) equations. The method of representing the system graphically through block diagrams lends itself much easier to such insights. This way of representing a system may be considered a graphical interface between the user and the equations describing the model. It is extensively used in the signal and systems/control engineering community, but not much in economics. It will be used in this thesis.

2.1 Phillips’ pioneering stock/flow model

In two seminal papers, Phillips (1954) and Phillips (1957) A.W. Phillips (who today is known in economics almost exclusively for something else – the *Phillips curve*) modeled the macroeconomy as a dynamic system consisting

of interconnected sub-entities and portrayed the model in the form of a block diagram¹. In the first paper he found the algebraic solutions to the models, while in connection with the latter paper he had access to an analog computer for numerical simulation. At the time, this was pioneering work. It was a second stage after his initial and physical hydraulic simulation macroeconomic model with vessels interconnected by tubes (Bissell; 2007).

Phillips had the advantage of a background as an electrical engineer, and acquainted himself with the fairly new discipline of control engineering and theory that had evolved strongly during the second world war. He saw that it could be applied to economics:

Recommendations for stabilizing aggregate production and employment have usually been derived from the analysis of multiplier models, using the method of comparative statics. This type of analysis does not provide a very firm basis for policy recommendations, for two reasons.

First, the time path of income, production and employment during the process of adjustment is not revealed. It is quite possible that certain types of policy may give rise to undesired fluctuations, or even cause a previously stable system to become unstable, although the final equilibrium position as shown by a static analysis appears to be quite satisfactory.

Second, the effects of variations in prices and interest rates cannot be dealt with adequately with the simple multiplier models which usually form the basis of the analysis. Phillips (1954), p. 290)

Figure 2.1 shows a facsimile from Phillips' 1954 paper, the simplest model with fixed prices. The model is in continuous time. The thick dotted line is added here. Above this line is the control subsystem. It compares actual output P to desired output P_d , and the error ε is fed into a PID controller which outputs a control variable called π . All these variables are time-varying money *flows*, with units² [$\$/y$]. The PID controller expresses the intervention strategy of the government. The control action – among other things a reaction to effects of an external disturbance u – is government

¹For readers not familiar with control system concepts, see appendix A for explanation of block diagram symbolism

²This is as opposed to what is typical in economic dynamics, where the time axis is partitioned into periods, and the denomination of *both* stock and flow variables becomes money *amount* [$\$$]. The dollar symbol will be used as a generic symbol for money in the following. And brackets [] signify denomination.

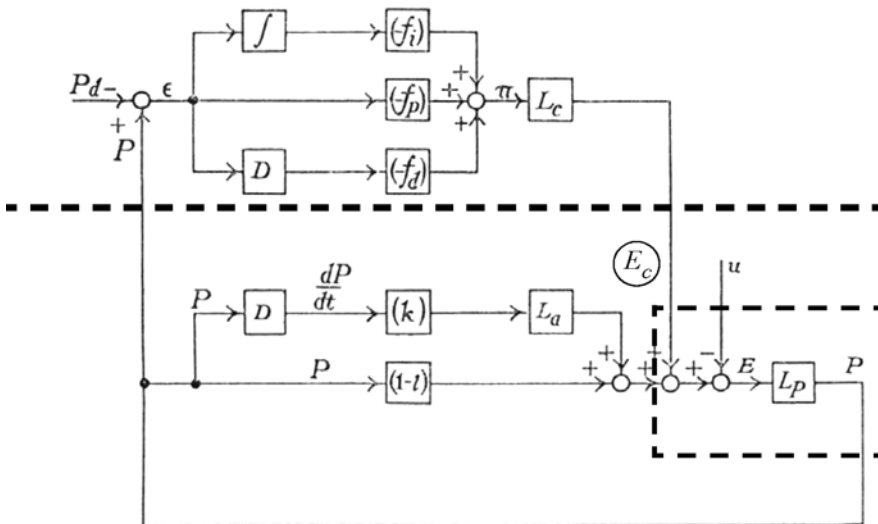


Figure 2.1: Facsimile of original Phillips block diagram

spending, indicated by the symbol E_c (a circle around the symbol indicates that it is inserted here and not part of Phillips' original figure). This spending action is not immediate, but somewhat lagged, indicated by the block L_c . Other lags are also indicated in the structure.

It is not the purpose here to discuss the control strategy proposed by Phillips, but only his model of the demand-to-output relationship, which is given by equation (2.1) below. The dotted rectangle (added here) indicates this part of his model, driven by the incoming aggregate demand flow E , resulting in output P . The block is called L_p in Phillips' notation. It is a *first order time lag*³, represented by the transfer function⁴

$$\frac{P}{E}(s) = \frac{1}{1 + T_p s} \tag{2.1}$$

which is equivalent to the linear first order differential equation

³We will also in between use the term "time constant", which is common in the controls systems community.

Also note at this stage that the term "time lag" is used in this thesis for something that is distinct from a *time delay*. This is discussed in section 2.4.

⁴See appendix A.

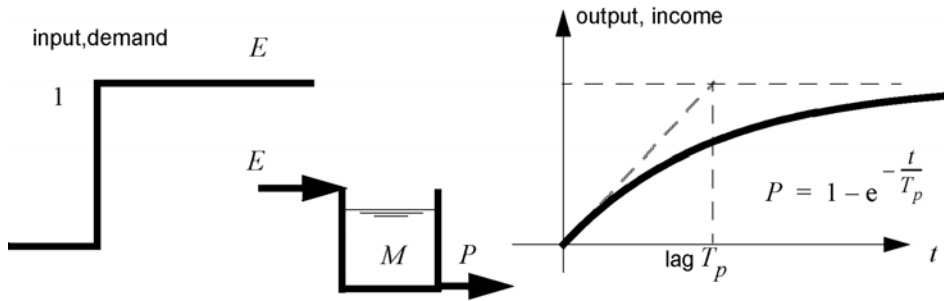


Figure 2.2: Step function and first-order time-lagged step response

$$T_p \dot{P}(t) = -P(t) + E(t) \quad (2.2)$$

The rationale for a model with the dynamics given by (2.1) (or (2.2)) has traditionally been explained as follows: The economy needs time to adjust to a change in demand. The first order time lag is the *simplest* continuous-time model for such dynamics. Thus Phillips and others (Godley and Cripps; 1983) use an “Occam’s razor” type of justification for their choice of model. In sections 2.8 and 2.9 we will strengthen the validity of the time lag model by deriving it from the fact that the economy may reasonably be assumed to consist of collections of a large number of agents (for instance aggregates of firms and of households).

The time lag model corresponds well (at least as a linearised approximation) to a buffer vessel: A sudden increase in the incoming flow will initially increase the level of fluid (in our case: money), which gradually leads to increased outflow. This is portrayed in figure 2.2. An economy modeled by (2.1) reacts to a jump in the demand flow with a time-dispersed outflow response asymptotically approaching the incoming flow level. When the outflow P (theoretically) has reached that asymptotic level, we have equilibrium. T_p is the time lag corresponding to the speed of adjustment.

2.2 Our approach

Inspired by Phillips’ simple first-order model of the macroeconomy, our approach may initially be explained by considering a subset, the aggregate of firms.

Assume a stepwise jump in demand. The increased inflow of money percolates through the interconnected network of tens of thousands of firms that constitute the demand-to-output part of the economy, and gradually (but not immediately) the effect will show up as increased income for workers and owners (i.e. the aggregate output flow). The lag for the firms sector is a consequence of *two different factors*: the time lag of each separate firm, and the degree to which the average unit of money flows via many other firms before it leaves the firm sector as income. We are especially interested in the effect of the last factor, and will present a related theorem in section 2.8.

Money stock $M(t)$ presently stored (or rather: in transit) within the block L_P must be the integrated difference between in- and outflows. Using the nomenclature in the Phillips block diagram, figure 2.1, we have the stock/flow balance equation

$$\dot{M}(t) = -P(t) + E(t) \quad (2.3)$$

At the same time we want a step response as in figure 2.2, corresponding to the transfer function (2.1). If we choose

$$P(t) = M(t)/T_P, \quad (2.4)$$

this is satisfied. Furthermore, equation (2.4) is intuitively appealing in the sense that the outgoing flow is proportional to money stock, which may then be regarded (via the physical “vessel” analogy) as a “pressure” – due to liquid level – driving this flow. And the larger the time lag T_P , the less flow P for a given M , i.e. a large time lag means that money has to accumulate significantly before it leads to increased outflow.

In Phillips’ model in figure 2.1, first order time lags are also used in two other places, labeled L_c and L_a . The first is a lag in the government’s control action (spending), which may be interpreted as either sluggishness in ascertaining the current economic situation, sluggishness in implementing the intervention policy, or a combination of both. The second lag accounts for some sluggishness in the investment response of investors to the rate of change in output (D is a differentiation operator in the block diagram), through an “accelerator” coefficient.

We note then, that first order time lags are used for two qualitatively different purposes in Phillips’ paper:

- To correctly describe the relationships between money stocks and flows – which may be called the “accounting part” of his model.

- To account for the fact that actions are not instantaneous – thus incorporating a *behavioural* characteristic.

This indicates a general usefulness of the first order time lag in economic modeling. We will employ it for both purposes in this thesis. We emphasise the distinction between the stock-flow accounting part of a model and its behavioural parts. The stock-flow accounting property is essential.

Until further notice we assume that the first order time lag is a realistic model of the monetary stock-flow dynamics of the aggregates of firms and households. We will also – in an average sense – consider this to be a satisfactory representation of input-output properties of *individual* firms and households. This will be justified later on.

In the following we will in between use the generic notation $F_i(t)$ and $F_o(t)$ for inflow and outflow, so that 2.3 becomes

$$\dot{M}(t) = -F_o(t) + F_i(t) \quad (2.5)$$

As mentioned earlier, money *stocks* in this thesis will be denominated with brackets, like this: [\$]. The dollar symbol is used as a generic symbol for money. Money *flows* will be denominated [\$/*y*], where *y* signifies "year".

All monetary variables are in nominal values. There will be hardly be any discussion of price or inflation dynamics, except fairly late in the thesis, in section 7.3.

2.3 Why time-continuous models?

Any model is only an approximation to the real phenomena it aims to represent. The actual economy runs in continuous time. But most dynamic macroeconomic models are time-discrete. Before the advent of computers and today's sophisticated simulation software, discrete-time models were easier to solve by hand and calculator. An early example of dynamical modeling with hand-calculated simulations is the multiplier-accelerator model as presented in Samuelson (1939). This history partly explains the discrete-time bias. It is also tempting to work in discrete time when important macroeconomic indicators are measured and made available with regular intervals, like years or quarters.

One problematic side effect of time-discrete models is that it becomes easy to confuse flows with stocks, since the denomination of flow entities – which actually is [\$/*y*] – in the discrete representation becomes [\$], just as if

it was a stock. As Michal Kalecki ironically remarked circa 1936, according to Joan Robinson:

I have found out what the science of economics is; it is the science of confusing stocks with flows (Robinson; 1982)

Keen (2009), pp. 163 – 164, makes another important point, that macroeconomic systems have a large span between the system time constants: There are slow-moving governments, firms and banks, and fast-moving households. This span is easily taken care of in a continuous-time-system framework. But in a discrete time representation one has to choose and stick with one specific discretisation step, which is an unsatisfactory compromise.

A further argument in favour of a continuous-time-system framework is that one may change system parameters without changing the simulation model. In a difference-equation framework (i.e. discrete-time model), most coefficients in the equations have to be changed whenever a system parameter is adjusted. The same has to be done if the time step (discretisation interval) chosen for the discrete model is changed.

In a continuous-time-system framework the (small) time step used in the numerical simulation algorithm – which is usually automatically set by the software and transparent to the user – may be adjusted without impacting any of the system's parameters, like time constants ("lags") and gains. The discrete vs continuous-time choice for modeling and simulation is considered very important. We will therefore dedicate all of chapter 3 to the topic.

A possible argument for time-discrete models might be that transactions between agents or sectors occur at discrete instances in time, and nothing happens in between. However, such an argument assumes regularly spaced events, while real-world transactions occur with *uneven* intervals. A precise and elegant way of accounting for such unevenly spaced events is using time-continuous models, but representing the discrete events with *impulse* (also called delta-, Dirac-) *functions*: If a unit of money is passed at time $t = t_1$ to an agent or a sector, this mathematically corresponds to an impulse function, commonly symbolised with $\delta(t - t_1)$. This function is a mathematical idealisation: it may be considered (one of several possible definitions) as the limit of a rectangular-shaped time function,

$$\delta(t) = \lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(t), \text{ with } \delta_\varepsilon(t) = \begin{cases} 1/\varepsilon, & |t| \leq \varepsilon/2 \\ 0, & |t| > \varepsilon/2 \end{cases} . \quad (2.6)$$

$\delta(t)$ has infinite amplitude and zero duration, but such that its area is unity. $\delta(t)$ is (as approximated by $\delta_\varepsilon(t)$) depicted to the left in figure 2.3. In an

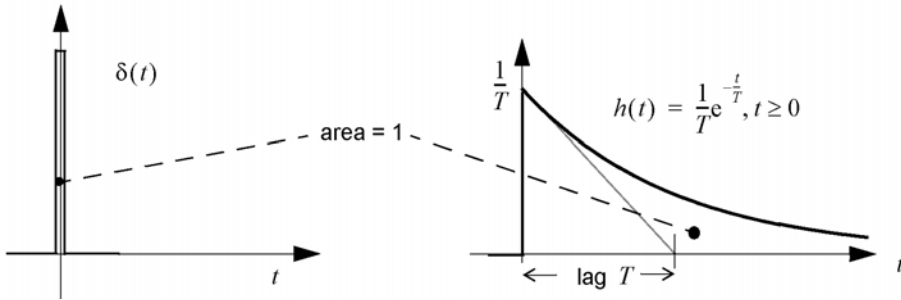


Figure 2.3: Impulse function (left) and impulse response (right)

economic model in continuous time, the impulse function allows a correct representation of time-discrete transactions: an amount of money Q passed to a sector or an agent at time t_1 is represented by the function $Q\delta(t - t_1)$. The denomination of this function is money *flow* [\$/y], while the area under the function has denomination money *amount* [\$]. The *impulse response* $h(t)$ of a unit (in our case an economic agent, a sector or the entire macroeconomic system) is defined as the output signal⁵ resulting from a 1 [\$] input at $t = 0$. The impulse response of a first order time lag is found by solving (2.5) with the input $F_i(t) = \delta(t)$. The result is

$$F_o(t) = h(t) = \begin{cases} \frac{1}{T}e^{-\frac{t}{T}}, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2.7)$$

$h(t)$ is shown to the right in figure 2.3. It is a *flow* [\$/y]. The area under $h(t)$ is unity. This is as expected, since money is neither created nor destroyed when passing by a unit. The mean time lag of $h(t)$ is

$$\int_0^{\infty} th(t)dt = \int_0^{\infty} t\frac{1}{T}e^{-\frac{t}{T}}dt = T \quad (2.8)$$

(The mean time lag T is also the value of t at the intersection between the tangent of $h(t)$ at $t = 0$ and the time axis, as indicated to the right in figure 2.3.)

⁵The symbol $h(t)$ is frequently reserved in the control systems (and signals and systems) literature to signify the output response to an impulse function, as distinct from responses to other input functions.

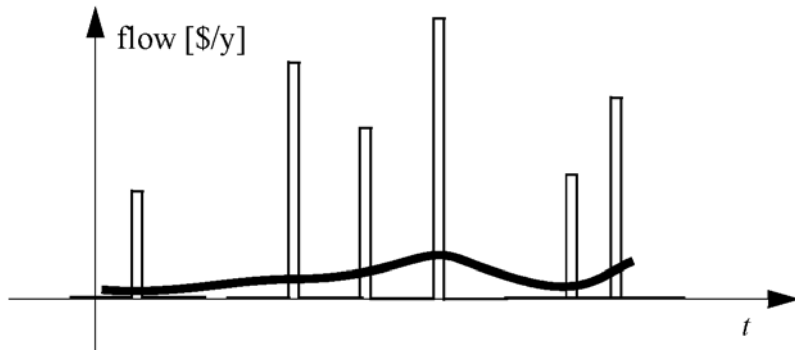


Figure 2.4: sum of areas under impulses = area under curve [\\$]

A further argument in favour of choosing the continuous-time framework is that a train of irregularly spaced impulses (which is the *precise* representation of transactions in continuous time) is very well approximated by a continuous flow *when the incidence of transactions is high*. This is portrayed in figure 2.4. Furthermore: When we are working with aggregates of many agents like firms and all households, “transaction impulses” between aggregates occur so frequently that continuous flow representation is quite satisfactory. The term “frequently” begs the question: “frequently, related to what?” The answer is that the dynamics of a sector with many units is sluggish, *related to the incidence of transactions*. The constant T_p in (2.1) above expresses this sluggishness (or “inertia”). In signals and systems terms we may say that the block L_p in Phillips’ model is a low pass filter with cutoff frequency $1/T_p$. Sharp fluctuations in the input will be smoothed out after having passed through. So the output will be similar whether the input is (faithfully) described as a chain of sharp spikes as shown in figure 2.4, or approximated by the corresponding smooth graph in the same figure.

(A similar argument is used for population and epidemiological models: infection events along the time axis, and also the current infected population size may be treated as continuous, even if both are discrete, Sterman (2000).)

2.4 Lags, delays, velocity and dispersion

In economics the term “lag” is used in discrete models to signify one discrete time step: a variable is a function of itself and some other variables from

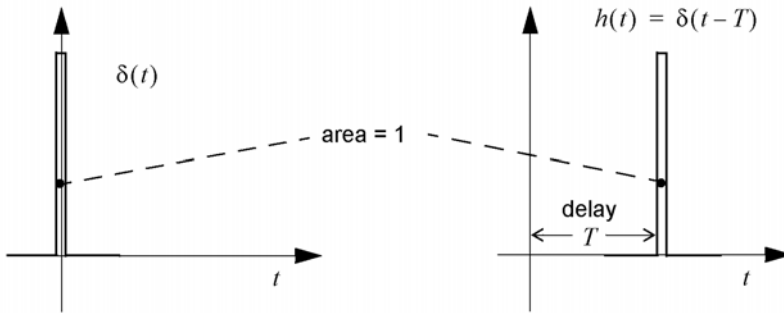


Figure 2.5: Impulse function and impulse response of a time delay unit

earlier periods. In a continuous-time setting this resembles what a control engineer would call a time *delay*. If the input is $\delta(t)$ to a unit having a time delay response, it will simply output a delayed impulse, as indicated in figure 2.5. This is different from the "vessel" model with impulse response as shown above in figure 2.3. In the following we will use the term "lag" for this type of response, and "delay" for the response in figure 2.5.

The delayed response corresponds to a unit (agent or sector) having the properties of a "tube", as opposed to the vessel analogy. We may think of the time delay in terms of a "packet" of money arriving at the tube's inlet, appearing at the outlet T time units later. Money stock $M(t)$ for the unit is then the "volume" = stock of money presently in the tube. For the special case with $F_i = F_o = \text{constant}$, M will also be constant. We have

$$M = FT, \text{ or } T = M/F \quad (2.9)$$

From the above follows that the unit's *velocity of money* is

$$v = 1/T \quad (2.10)$$

Velocity v [$1/y$] is a central concept in macroeconomics, and will be used frequently throughout⁶.

Consider the case when the unit is a model of an entire economy with no government, no savings and where investment stems from profits that are paid to owner households along with wages to workers (no lending from banks). Then the "tube" metaphor for the macroeconomy may be portrayed as in figure 2.6. The delay T is indicated along the tube's length since

⁶A better term for v might be transaction frequency. It will also be used in between.

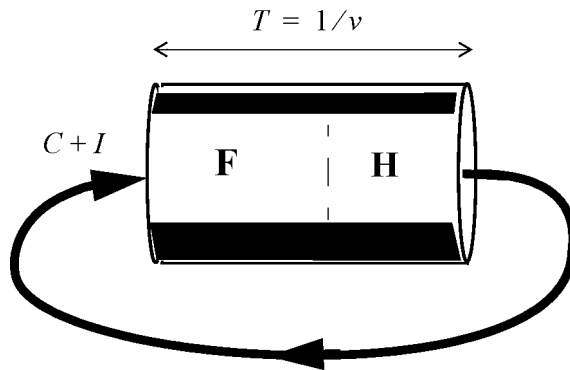


Figure 2.6: A "tube" with firms (F) and households (H)

it is proportional to it. Every round for any “packet” of money will in accordance with the tube metaphor lasts T [y] time units. Stated another way, a “packet” will traverse $v = 1/T$ [$1/y$] rounds during one time unit.

The delay associated with flows in general (as in process plants, pipelines, etc.), will in the case of money be the time a given amount spends between arrival and departure at a given unit (here: the entire aggregate of firms and households), while flows between units are instantaneous. *Money always resides with some unit.*

2.4.1 Dispersion in time

The introduction of money velocity above was based on the very restrictive assumption of constant flow. The tube metaphor also presupposes that all parts of money stock circulate at the same speed. These assumptions will not be needed with the superior vessel metaphor. If we compare figures 2.3 and 2.5, we note that both responses have the same area and the same mean time lag (or delay), T . The important difference is that the time lag response in 2.3 is *dispersed in time*. This has an intuitive appeal: if an amount of money is received by some sector at some instance, this amount will be spread out in time on its way through. Parts of it will follow a very convoluted path in the sense that it will be used by many agents for transactions *within* the sector, before being spent *out of* the sector.

The same holds for money being received by a single agent⁷ within a

⁷In this thesis the term "agent" signifies the smallest component in the macroeconomy,

sector at a certain moment (example: a monthly wage); it will not all be spent at once but spread out over time. The time lag (vessel) model expresses the dispersed character of the response in a reasonable manner, while the time-delayed (tube) response does not give dispersion at all.

The dispersion-in-time property, which holds for all input-output relationships for agents and sectors, invalidates the approach of analysing monetary circuit dynamics by assuming that these unfold in distinct and concluded "periods". This common simplifying assumption among economists, both in the mainstream and among those belonging to the Post Keynesian and Circuitist schools, will be critiqued in section 3.1.

The time lag model has a time-dispersed, not exactly delayed response. But it still gives the *average* time money takes to complete an economic circuit (it also has an inverse; average money velocity). Equation (2.9), which was based on the the tube metaphor with $F_i = F_o = \text{constant}$, also holds for the vessel model, as long as we confine ourselves to the constant flow case⁸.

For a first order time lag unit we assume, as already stated with (2.4) and (2.3):

$$\dot{M}(t) = -F_o(t) + F_i(t) \quad (2.11a)$$

$$F_o(t) = M(t)/T \quad (2.11b)$$

This model allows varying flows, and differences between in- and outflows, even with $v = 1/T = \text{constant}$. And if we consider money velocity T to be a behavioural characteristic for the unit, it is reasonable that it relates only to outflow, not inflow, since the unit only has control over its outflow. We will later on use the term *discretionary* for an outflow that is controlled by the agent, and *non-discretionary* for flows demanded from the agent but normally outside the agent's control (debt service paid by a debtor is one example, tax payment is another).

for instance a household or a firm. We also sometimes use the terms "unit" or "node" for this. The last term is appropriate because it indicates that we describe the macroeconomy as a network of nodes, interconnected by money flowing between them.

⁸A fascinating result is called *Little's law* after the discoverer John Little (Serman; 2000), pp. 421 - 423. When inflows and outflows are constant, *any* linear input-output unit – regardless of impulse response shape as long as the area under the response is unity – satisfies 2.9 (for a more precise statement, see subsection 2.6.2).

One may also make a chain of (not necessarily identical units), and the T for the whole chain will be the sum of the individual lags, and satisfy 2.9 with an M being the sum of the individual M 's "in transit" with the respective units.

2.4.2 Impulse response vs step response

If we compare figures 2.2 and 2.3, we observe how two different input signals (step function and impulse function) give different output responses but with a common characteristic; the shape of the response is a decaying exponential with the same time constant. The first response is called a (unit) step response. The step response is the time integral of the impulse response, and the unit step input function is the integral of the impulse (input) function. Both responses can and will be used to equivalently characterise a given unit. We note that while the impulse represents a one-shot payment of money, the step function represents a persistent *flow* of money starting at time $t = 0$.

In chapter 4 we will use the impulse response concept as a convenient way to represent debt service: a loan given at time $t = 0$ is represented by an impulse function and generates a debt service flow which then may be considered to be the impulse response of a "debt service subsystem". This, together with modeling the debtor part of the economy as a first order time lag unit, enables a simple but still valid analysis of the results of persistent recycling of debt service income into new and overlapping loans, something that is usually abstracted from in the monetary circuit literature partly because it has been considered too difficult (due to time dispersion of money flows).

2.4.3 Agent dynamics with no inflow

We have until now considered the dynamics of units with a one-shot payment or a constant inflow of money starting at $t = 0$. If we alternatively consider a situation with a certain initial money stock but no inflow, i.e. $F_i(t) = 0$, then our agent, from solving (2.11), spends her money following a decaying exponential curve, which is reasonable behaviour in a situation with zero inflow. See figure 2.7. This gives further support to the first order time lag model.

2.5 On saving, "hoarding" and liquidity preference

In economic textbooks one mostly confines the description of what agents do with their money to the alternatives "spending" and "saving". If an income flow to an agent is called F , spending is given by $(1 - \sigma)F$ and saving by σF , where σ is "the propensity to save", $\sigma < 1$ (we use the symbol σ throughout for the savings coefficient, since we reserve s as symbol for the derivative operator/Laplace transform variable). More sophisticated models

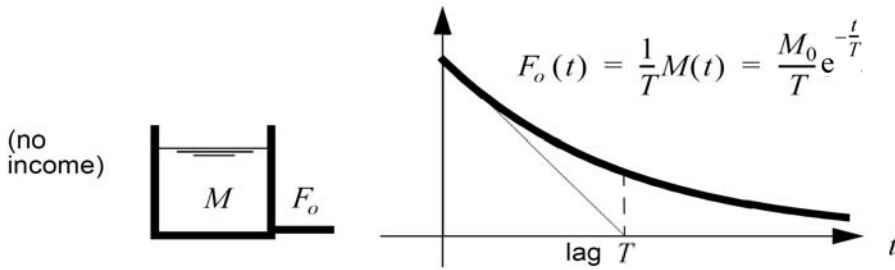


Figure 2.7: Spending with no income and $M(t_0) = M_0$

incorporate a "portfolio" for what sort of saving the agent does; one may for instance distinguish between stocks and bonds in such a savings portfolio.

Some Post Keynesian economists, among them Giuseppe Fontana (2000), pp. 34 – 36, makes the point that one additionally should distinguish between saving and *hoarding*. The latter is holding on to money, while the former is converting one's money to some financial instrument. This is a good point. But they go wrong by accounting for hoarding with a "*marginal*⁹ propensity to hoard", which called $1 - m$, with $m < 1$. The share for buying securities out of income, is then m .

For an income flow F_i , then

$$\text{spending} = (1 - \sigma)F_i, \text{ saving} = \sigma m F_i, \text{ and hoarding} = \sigma(1 - m)F_i \quad (2.12)$$

The argument against this representation may most conveniently be done using block diagrams. See figure 2.8. The upper diagram depicts Fontana's and others' structure with a propensity to hoard, while the lower contains the block diagram of the first order time lag which will be used frequently throughout this thesis.

The upper structure implies that a unit receiving a money flow immediately channels this onwards, either to buy some financial instrument (bonds and stocks), to consumption – *or to the unit's money hoard*. This has two problematic implications: the money hoard will grow persistently over time without bounds, and the hoard is not used by the unit as a buffer in the circulatory system, but removed from circulation. The first is obviously impossible, the second contradicts the actual behaviour of units like households and firms.

Augusto Graziani writes:

⁹"Marginal" is in Fontana's paper used in a non-marginal sense. $1 - m$ is simply the share of income flow that is split off from other spending.

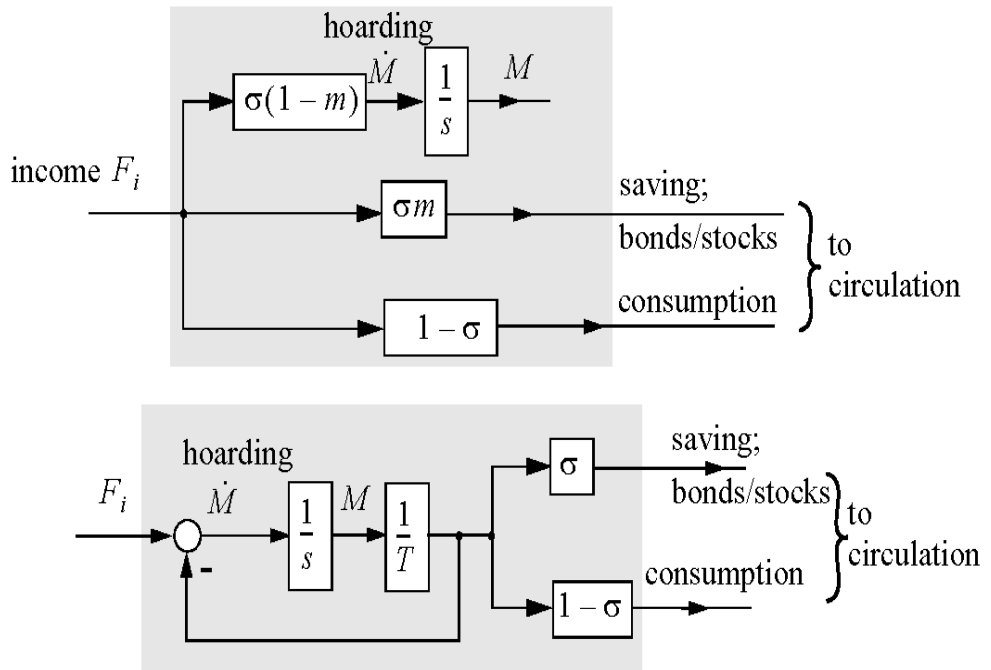


Figure 2.8: Two structures with hoarding

To the extent that wage earners, instead of spending their whole incomes either on commodities or securities, keep their savings in bank deposits, the firms are unable to repay their banks debts. If the previous level of output has to be preserved, the firms have to get new loans from the banks, which means that the 'stock' of money has increased (Graziani; 1996), p. 144.

He is correct that loans must be given to compensate for existing ones being re-paid, if circulation is to continue. But as long as this happens, output may be upheld with zero net lending (i.e. no increase in money stock) if wage earners do what we might call *in-flow hoarding*, as indicated in the lower half of figure 2.8. We will return to this topic in subsection 2.7.5.

When money is spent in the in-flow-hoarding model, one may account for the other portfolio choices by splitting the outflow via the remaining propensities (the coefficients) that is used in the figure. The propensity to hoard model has three outputs decided by two coefficients, m and σ . While

our model has a time lag and two outputs decided by one coefficient, σ . In our model the unit adjusts the hoarding rate by changing T in the same direction.

A comprehensive critique of the concept of a propensity hoard is done in (Andresen; 2006).

If one mainly considers a unit's cash or deposit account (the "hoard") as a buffer between in- and outflows – as in this thesis – the time lag (inverse money velocity) also emerges out of the model. This parameter gives a representation of *liquidity preference*. The concept is defined in Keynes (1936), ch. 13, part II. Keynes introduced the term to characterise how agents adjusted their spending behaviour in reaction to a change in the interest rate, and as a precaution against possible extraordinary future expenses. We will later on, in subsection 6.2.3, use his second point in a similar but slightly different way, to characterise agents' mood change in an economic downturn. And money velocity (or its inverse, the time lag T) in the lower model expresses this well. In a downturn scenario with increasing pessimism, agents will generally hold back in their spending, trying to safeguard their "money hoard". Increased liquidity preference corresponds to a reduced v (increased T) by the agent. *Money velocity is a behavioural variable*. Metaphorically, decreasing v means that the agent (household, firm) turns down the outflow "tap" on his/her "vessel". When many units do this, the result is a general decrease in consumption and spending, in the next round increasing pessimism, leading to even more liquidity preference and worsening the crisis. This can lead to a death spiral towards depression. We will model and simulate such a process in chapter 6.

2.6 Time lags for interconnected units

This section develops and discusses some rules for how to simplify interconnected time lag units. The goal is to find a reasonable single unit approximation, in the sense that the time lag of this single unit reasonably approximates the time lag for the network. We accept that the precise dynamics of the network will remain unknown – *we confine ourselves to finding a fairly equivalent time lag*. This is a special type of model reduction technique. Model reduction is frequently used in the control systems community. The discussion to follow is valid for linear systems as such, not only economic systems.

2.6.1 Micro to macro?

Before proceeding however, some remarks about aggregation and models of unit behaviour. Models of an individual agent used in neoclassical economics are often assumed to also represent the *aggregate* of the agents in the macroeconomy. In such models households maximize intertemporal "utility" subject to their budget constraint. Robert Solow is very critical:

... a school of macroeconomic thought that dominates many of the leading university departments and some of the best journals, not to mention the Federal Reserve Bank of Minneapolis. They mean a macroeconomics that is deduced from a model in which a single immortal consumer-worker-owner maximizes a perfectly conventional time-additive utility function over an infinite horizon, under perfect foresight or rational expectations, ... (Solow; 2008)

Our household unit is a much simpler creature, (s)he just tries to uphold a buffer of money with respect to his/her own situation and how (s)he sees the general economic prospects. Later on, we will distinguish between "worker households" who spend all they receive – and "capitalist households", who in addition to upholding a buffer of money and consuming, also extend loans and do real-economy investment.

There is a comprehensive literature in economics about whether one reasonably may construct macroeconomic (aggregated) models based on some replication of assumed models of individual agent behaviour. One central point of disagreement is whether a large collection of heterogenous individuals who obey the "Law of Demand" (downward sloping demand curves) imply an aggregate (market) which obeys the same law. Keen (2011) argues against this, among other things referring to the Sonnenschein-Mantel-Debreu conditions which proves that this is generally not the case.

A paper arguing against the proposition that a model of "the whole" (in physics and other fields, not only economics) can be inferred from properties of its parts, is (Anderson; 1972). He writes:

The main fallacy in this kind of thinking is that the reductionist hypothesis does not by any means imply a "constructionist" one: The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. In fact, the more the elementary particle physicists tell us about the nature of the fundamental laws, the

less relevance they seem to have to the very real problems of the rest of science, much less to those of society. The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other (p. 393).

While agreeing wholly with this, the approach of this thesis is to *circumvent* the "microfoundations for macro" controversies in economics by looking at both individual agents and sectors in a different and much simpler way: they are all modeled as stock-flow units connected via monetary flows. A unit is characterised by one main parameter, its time lag. And with *this* approach, it will be argued that it becomes meaningful to use the properties of individual units to construct aggregate models.

Some might object that this is too simple to allow any non-trivial and useful insights to emerge. The reader must be the judge of this.

2.6.2 Definitions and introductory remarks

Until now all units have been assumed to have first order dynamics of the type in figure 2.2, which is quite restrictive. We will from now rescind this assumption, and allow all types of units that satisfy the following criteria for their impulse responses:

$$\text{causality, i.e. } h(t) = 0, t < 0 \quad (2.13a)$$

$$\text{non-negative: } h(t) \geq 0, \forall t > 0 \quad (2.13b)$$

$$\text{unit area: } \int_0^{\infty} h(t) dt = 1 \quad (2.13c)$$

$$\text{non-zero duration: } h(t) > 0 \text{ for some finite time interval} \quad (2.13d)$$

We will call the class of impulse responses satisfying (2.13) "*PUA*" (positive, unit area). For money stock-flow applications, these criteria are obviously fulfilled: money cannot emerge from a unit before it is received (2.13a), a money outflow cannot be negative (2.13b), money is not created or de-

destroyed¹⁰ by the unit (2.13c), and money received is not passed on in one lump (2.13d)(this last condition may be violated for certain units in certain cases, but such non-dispersion cases are assumed to occur so rarely that they may be abstracted from).

For an arbitrary type of PUA unit, we now define the mean time lag \bar{T} as the position along the time axis of the centroid of the area under the impulse response $h(t)$. Introducing the Laplace transform $h(s)$ (= the transfer function)¹¹ of $h(t)$ for a unit, we have that the mean time lag is

$$\bar{T} = \int_0^{\infty} th(t)dt = \int_0^{\infty} th(t)e^{-st}dt \Big|_{s=0} = -\frac{d}{ds}[h(s)] \Big|_{s=0} = -h'(0) \quad (2.14)$$

Note the compact notation of the rightmost term, which will be used for convenience in the following. Dependence on s , not t , is implied. In the more general case, if the response in question does not have unit area, we have

$$\bar{T} = \frac{\int_0^{\infty} th(t)dt}{\int_0^{\infty} h(t)dt} = \frac{\int_0^{\infty} th(t)e^{-st}dt}{\int_0^{\infty} h(t)e^{-st}dt} \Big|_{s=0} = \frac{[-\frac{d}{ds}h(s)]_{s=0}}{h(s)|_{s=0}} = \frac{-h'(0)}{h(0)} \quad (2.15)$$

We note that unit area implies $h(s)|_{s=0} = 1$ ¹².

We introduce a class of PUA units which have a rational transfer function

$$h_r(s) = \frac{n(s)}{d(s)} = \frac{1 + b_1s + \dots + b_{n-1}s^{n-1}}{1 + a_1s + \dots + a_ns^n}, \quad (2.16)$$

¹⁰The sole exception is when the unit is a commercial bank, or the commercial bank sector as an aggregate. This is first discussed in chapter 5.

¹¹For convenience we use $h(\cdot)$ both in the time domain and the s domain, in spite of these functions being mathematically different. The context will tell which interpretation applies. $h(s)$ then signifies the Laplace transform of $h(t)$.

¹²In control systems jargon $h(s)|_{s=0}$ is called the *static gain* of the transfer function $h(s)$. A constant stepwise input from $t \geq 0$ onwards may be interpreted as a cosine function with zero frequency, switched on at $t = 0$ – thus the adjective "static". With static gain = 1, the resulting output becomes constant with the same amplitude as the input, after some transient period.

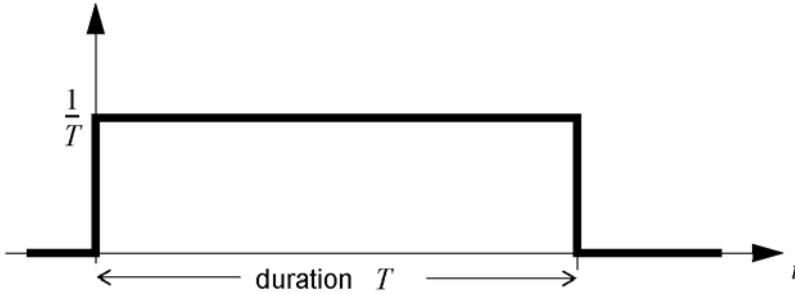


Figure 2.9: Rectangular PUA pulse

(Note that h_r doesn't have to be *strictly proper*: the degree of the numerator may be equal to the degree of the denominator. If that is the case, the unit in question has an instantaneous "by-pass" connection from input to output. In monetary terms, this means that this unit *immediately* spends some of the money received. This special case is not considered to be important, and thus not considered further in this thesis.)

Using (2.14), we get

$$\bar{T} = \frac{n(0)d'(0) - d(0)n'(0)}{d^2(0)} = \frac{d'(0) - n'(0)}{1} = a_1 - b_1 \quad (2.17)$$

For the special and simplest case of the first order time lag with $n(s) = 1$ and $d(s) = 1 + Ts$, we know that $\bar{T} = T$, which is confirmed by (2.17). For higher-order PUA rational transfer functions we note that \bar{T} is independent of the coefficients b_i and a_i , $i > 1$.

Let us now consider a unit with a *non-rational* transfer function

$$h_{nr}(s) = \frac{1}{Ts}(1 - e^{-Ts}), \quad (2.18)$$

the corresponding impulse response $h_{nr}(t)$ is PUA with a rectangular shape as shown in figure 2.9. By visual inspection the mean time lag is obviously $\bar{T} = T/2$. It may also be found via (2.14):

$$\bar{T} = \lim_{s \rightarrow 0} \left(-\frac{\partial}{\partial s} \left\{ \frac{1 - e^{-Ts}}{Ts} \right\} \right) = \frac{T}{2} \quad (2.19)$$

2.6.3 The mean time lag for units in series and in parallel

For N PUA transfer functions in a series connection, the resulting transfer function will also be PUA. (2.14) gives

$$\begin{aligned}\bar{T} &= -h'(0) = -(h_1 h_2 \cdots h_N)'(0) = -(h_1'(0) + h_2'(0) + \cdots + h_N'(0)) \\ &= T_1 + T_2 + \cdots + T_N\end{aligned}\quad (2.20)$$

This is reasonable, considering that \bar{T} may be seen as the lag of a flow which is transmitted through a chain of N “agents” (referring to our economic application). It also agrees with Little’s law, mentioned earlier (Sterman; 2000), pp. 421 - 423. The impulse response of a series connection of subsystems as in (2.20) is the convolution of PUA impulse responses in the time domain:

$$h(t) = h_1(t) * h_2(t) * \cdots * h_N(t) \quad (2.21)$$

An example: if we make a serial connection of three identical units of the type (2.18), the transfer function is

$$h_{nr}^3(s) = \left(\frac{1 - e^{-Ts}}{Ts} \right)^3 \quad (2.22)$$

From (2.14), the mean time lag is

$$\bar{T} = \lim_{s \rightarrow 0} \left(-\frac{\partial}{\partial s} \left\{ \left[\frac{1 - e^{-Ts}}{Ts} \right]^3 \right\} \right) = \frac{3T}{2}, \quad (2.23)$$

which can be more easily found by inspection of the resulting PUA impulse

response $h_{nr}^3(t)$, which is the triple convolution of the rectangular PUA pulse with itself. See figure 2.10. The response $h_{nr}^3(t)$ consists of three smoothly spliced parabolic sections, with common tangents at the splicing points. This response is a soft smooth pulse where nothing happens immediately (as opposed to the response of the first order time lag unit) and which tapers to zero after a finite time (as opposed to the infinite-tail response of any PUA unit with a rational transfer function).

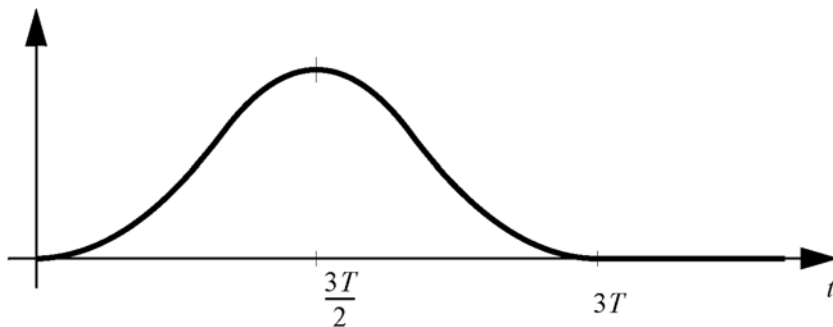


Figure 2.10: PUA impulse response $h_{nr}^3(t)$

A “limit impulse response” for serially connected units has a normal distribution shape

(First, a warning: This result does not have any use for later considerations and applications. I still insert it here since I find it very fascinating and haven't found it in the literature ...)

A PUA response is mathematically similar to a probability distribution function in the sense that both are PUA (except that a p.d.f.-shaped time function doesn't necessarily satisfy condition (2.13a)). From probability theory we know that if we convolve N p.d.f.'s, the result is the p.d.f. for the sum of the respective random variables, assuming they are i.i.d., see for instance Casella and Berger (2002), p. 215. And the Central Limit Theorem (ibid., p. 236) tells us that the resulting p.d.f. will tend to a normal distribution when N is large. If we return to the time domain and apply this to a linear dynamic system consisting of a large number of N serially connected similar PUA subsystems, this means that the impulse response of the total system will have a shape approaching that of the normal distribution, with mean time lag corresponding to the mean value of the distribution. To explore this, we introduce the PUA transfer function

$$h_N(s) = \frac{1}{\left(1 + \frac{\bar{T}}{N}s\right)^N} \quad (2.24)$$

We use it in for a chain of N serially connected identical first order time lags. Since each lag in the chain is \bar{T}/N , rule (2.20) tells us that the system's mean lag will be invariant $= \bar{T}$.

Incidentally, we note that

$$\lim_{N \rightarrow \infty} h_N(s) = e^{-\bar{T}s} \quad (2.25)$$

so that the impulse response in the limit is trivially an impulse delayed by \bar{T} , as in figure 2.5. We want, however, to examine the responses for large but finite N . The impulse response corresponding to (2.24) is

$$h_N(t) = \frac{1}{(N-1)!} \left(\frac{N}{\bar{T}} \right)^N t^{N-1} e^{-N \frac{t}{\bar{T}}} \quad (2.26)$$

Figure 2.10 shows a selection of responses, with $N = 1, \dots, 100$ and $\bar{T} = 1$. Responses for $N = 1, 2, 3$ are emphasised. The corresponding normal distribution shape has mean \bar{T} and variance \bar{T}^2/N , and is shown as a dotted graph. These responses are all from transfer functions with a constant = 1 in the numerator (no zeros). (We will introduce a zero and use $N = 2$ for Monte Carlo simulations of an economic network in section 2.9. We will see that a zero and two poles are sufficient to generate a very wide range of different-shaped responses.)

Units in parallel

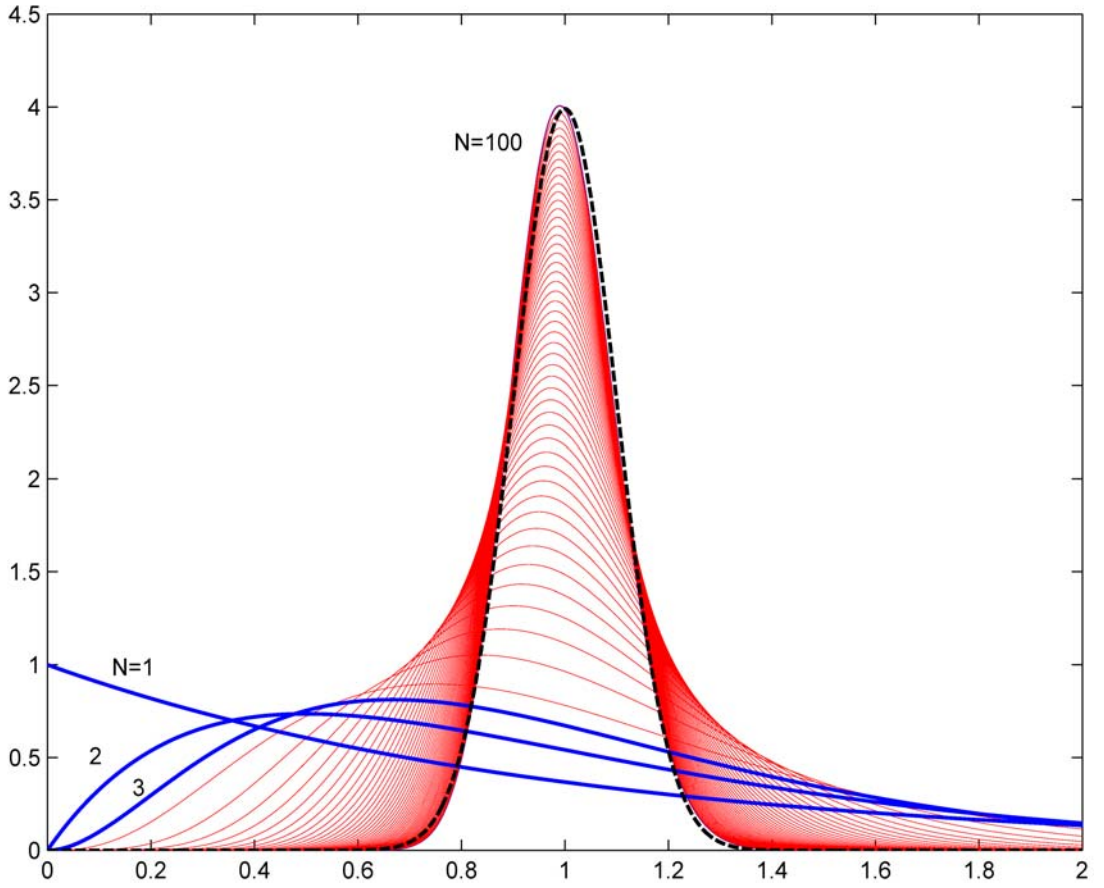
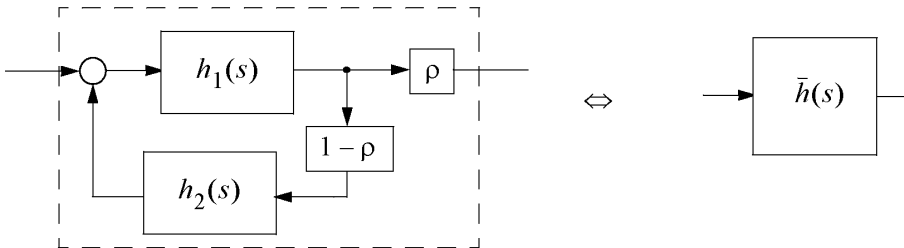
Now to the case of N arbitrary PUA units with time lags T_i , connected in parallel. The transfer function becomes

$$h = \mu_1 h_1 + \mu_2 h_2 + \dots + \mu_N h_N \quad (2.27)$$

Here we require all constant coefficients $\mu_i > 0$, and $\mu_1 + \mu_2 + \dots + \mu_N = 1$ so that $h(s)$ will also be PUA. The time lag for $h(s)$ is then a weighted average,

$$\begin{aligned} \bar{T} &= -h'(0) = -(\mu_1 h_1 + \mu_2 h_2 + \dots + \mu_N h_N)'(0) \\ &= \mu_1 T_1 + \mu_2 T_2 + \dots + \mu_N T_N \end{aligned} \quad (2.28)$$

The series (2.20) and parallel (2.28) rules for calculating time lags may be then employed to find the mean the time lag of a network of a more arbitrary structure, which we will do later.

Figure 2.11: Impulse responses $h_N(t)$ Figure 2.12: Equivalent transfer function $\bar{h}(s)$ from feedback

2.6.4 Mean time lag for a system with feedback

Consider the block diagram in figure 2.12. Here $h_1(s)$ and $h_2(s)$ are transfer functions of arbitrary PUA units, with time lags T_1 and T_2 . Coefficients ρ and $1 - \rho$, with $0 < \rho < 1$, are necessary to ensure that the resulting unit with transfer function $\bar{h}(s)$ is PUA. We have

$$\bar{h} = \frac{\rho h_1}{1 - (1 - \rho)h_1 h_2} \quad (2.29)$$

Using (2.29) in (2.14) and partial differentiation, the mean time lag of $\bar{h}(s)$ is

$$\begin{aligned} \bar{T} &= -\frac{\partial}{\partial s} \left\{ \frac{\rho h_1(s)}{1 - (1 - \rho)h_1(s)h_2(s)} \right\} \Big|_{s=0} \\ &= \frac{\partial}{\partial h_1} \left\{ \frac{\rho h_1}{1 - (1 - \rho)h_1 h_2} \right\} T_1 + \frac{\partial}{\partial h_2} \left\{ \frac{\rho h_1}{1 - (1 - \rho)h_1 h_2} \right\} T_2 \Big|_{h_1=1, h_2=1} \\ &= \frac{T_1 + (1 - \rho)T_2}{\rho} \end{aligned} \quad (2.30)$$

We may relax the condition that $\bar{h}(s)$ shall be PUA. The gain at the output in figure 2.12 may differ from ρ and have an arbitrary value. This will not change the lag \bar{T} of $\bar{h}(s)$, only its static gain.

We will use the feedback rule in subsection 3.1.1.

2.6.5 Mean time lag from a state space representation

Finally, consider a linear SISO (single-input, single-output) system in state space form with a positive – not necessarily unit area – impulse response:

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}u \quad (2.31a)$$

$$y = \mathbf{c}^T \mathbf{x} \quad (2.31b)$$

The transfer function is

$$h(s) = \mathbf{c}^T (sI - A)^{-1} \mathbf{b} \quad (2.32)$$

From (2.15) and (2.32) we then have the mean time lag

$$\bar{T} = \frac{-h'(0)}{h(0)} = \frac{\mathbf{c}^T [-(sI - A)^{-1} I (sI - A)^{-1}]_{s=0} \mathbf{b}}{-\mathbf{c}^T A^{-1} \mathbf{b}} = -\frac{\mathbf{c}^T A^{-2} \mathbf{b}}{\mathbf{c}^T A^{-1} \mathbf{b}} \quad (2.33)$$

The state space representation (2.31) will be used for Monte Carlo simulations of a network with 150 agents in section 2.9.

2.6.6 Mean time lag via the step response

All time lag rules until now have been explored based on a PUA system's impulse response. We will now instead define the lag based on a PUA system's *unit step* response. For convenience we repeat these two types of responses, for the 1st order time lag system, $h(s) = 1/(1 + Ts)$. See figure 2.13. Before proceeding, note that in the case $h(s) = 1/(1 + Ts)$, the two

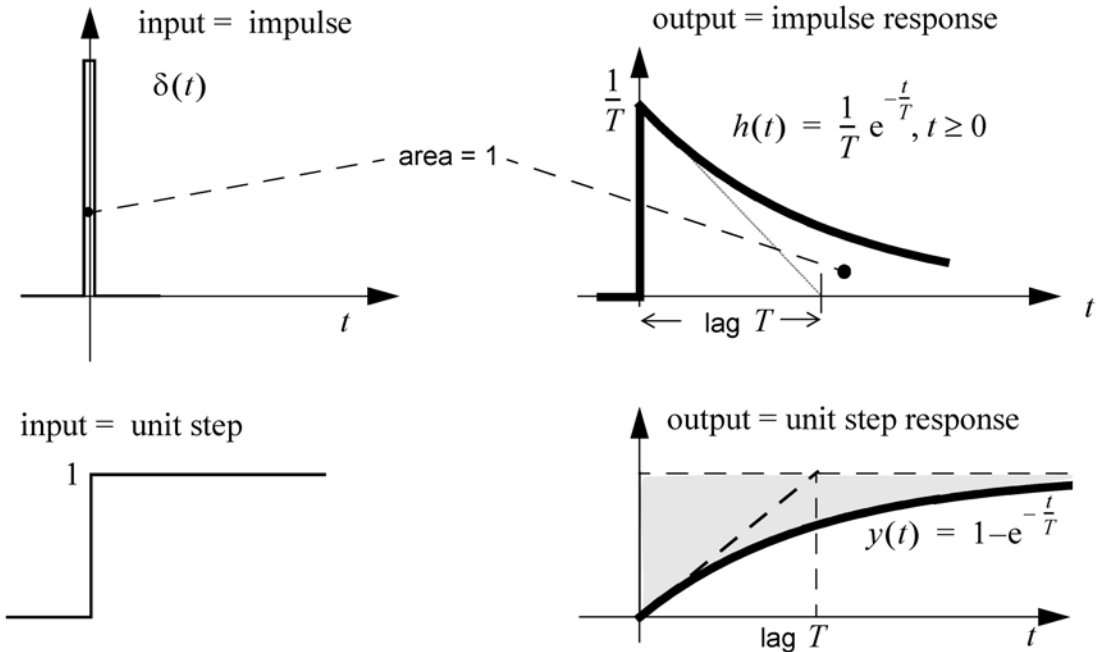


Figure 2.13: Impulse and unit step response for $h(s) = 1/(1 + Ts)$

responses have similar (decaying exponential) shapes and the step response starts at the origin; this is always the case if the unit has a strictly proper¹³ transfer function. But the derivation to follow will be valid for any PUA system of the type (2.16), repeated here for convenience:

$$h(s) = \frac{n(s)}{d(s)} = \frac{1 + b_1s + \dots + b_{n-1}s^{n-1}}{1 + a_1s + \dots + a_ns^n} = \frac{F_o(s)}{F_i(s)} \quad (2.34)$$

¹³Strictly proper means that the degree of the polynomial in s in the denominator is higher than the degree in the numerator.

where F_i and F_o as earlier are the generic input and output symbols. For any such system the unit step response will increase monotonically and converge towards 1. With a persistent constant input (the unit step flow), the system converges to an equilibrium where inflow F_i and outflow F_o is equal $= F$, and where the buffer stock M becomes constant. M must be the time-integrated difference between in- and outflow in the initial build-up phase, since we have $\dot{M} = F_i - F_o$. In figure 2.13 M is the shaded area. We define the mean lag such that $F_o = F = M/\bar{T}$, as already indicated in (2.11). We will now use the Final Value Theorem for Laplace transforms, which says that for a time-dependent function known to be tending to a constant value as $t \rightarrow \infty$, we have

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} s f(s) \quad (2.35)$$

We then get for the equilibrium value of M :

$$\begin{aligned} \lim_{t \rightarrow \infty} M(t) &= \lim_{s \rightarrow 0} s M(s) = \lim_{s \rightarrow 0} \left(s \frac{M}{F_i}(s) F_i(s) \right) = \lim_{s \rightarrow 0} \left(s \frac{M}{F_i}(s) \frac{1}{s} \right) \\ &= \lim_{s \rightarrow 0} \left(\frac{1}{s} \frac{F_i(s) - F_o(s)}{F_i(s)} \right) = \lim_{s \rightarrow 0} \left(\frac{1}{s} \left[1 - \frac{F_o(s)}{F_i(s)} \right] \right) \\ &= \lim_{s \rightarrow 0} \left(\frac{1 - h(s)}{s} \right) \\ &= \lim_{s \rightarrow 0} \left(\frac{a_1 s + \dots + a_n s^n - b_1 s - \dots - b_{n-1} s^n}{s(1 + a_1 s + \dots + a_n s^n)} \right) \\ &= a_1 - b_1 \end{aligned} \quad (2.36)$$

Since the equilibrium value of F is unity, this gives

$$\bar{T} = \frac{M}{F}(s \rightarrow 0) = a_1 - b_1 \quad (2.37)$$

Thus we get the same time lag based on this definition as via the earlier definition which was based on the impulse response, (2.14). If the input step is of amplitude $A \neq 1$, the area M will be scaled by the same factor (as this is a linear system). We get $\bar{T} = \frac{M}{F} = A(a_1 - b_1)/A = a_1 - b_1$; the same lag.

We may formulate the result as a verbal rule:

"regardless of the shape of the PUA system step response, the time lag for the system is the equilibrium buffer (indicated by the area between the input and output flow graphs) divided by the height of the input step (which is the size of the equilibrium in- and outflow)".

For the discrete time case this rule is described in (Godley and Cripps; 1983) and (Godley and Lavoie; 2007).

Sterman (2000) discusses this for the continuous-time case. But instead of characterising a unit through its (deterministic) impulse response, Sterman (on p. 423) employs the probability density function for when a "particle" appears in the outflow when it has been input at $t = 0$. The result however, is the same whether one employs a p.d.f. or an impulse response.

We use deterministic dynamics in this thesis due to the homogeneity of money: all "particles" in a monetary flow are identical. We may therefore think of a money flow as if it was a homogenous liquid. This as opposed to a network where the "particles" are heterogenous, for instance like the letters and parcels in a postal system, used as an example in Sterman's book. There one should work with probabilities for when a letter or parcel will arrive.

2.7 Application: an economy with households and firms

Before proceeding with further technical development of the "network of time lags" concept, we will in this subsection apply some of what has been derived until now. We will consider a simple textbook variant with households and firms, with no government and no financial sector. The diagram in figure 2.14 is a "monetary-flow-and-vessels diagram" representation of this economy. Below is the (mathematical) block diagram of the same system. In the figures we have these entities:

Y_d = aggregate demand [$\$/y$].

Y_o = aggregate output [$\$/y$].

$\pi = 1 - w$ = profit share of output [], w is workers' share.

Π = profit [$\$/y$]; all profits are invested, and there is no external source of investment at this stage.

W = wages [$\$/y$]; all wages are consumed, and there is no household savings sink (or borrowing source).

C = consumption money stocks in the two sectors are M_F and M_H .

I = investment [$\$/y$].

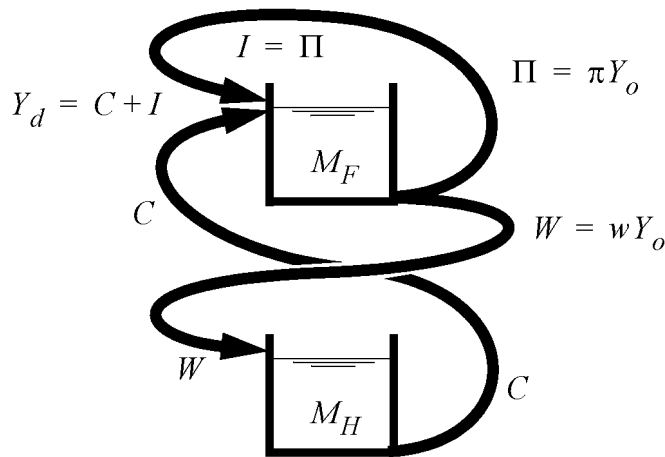


Figure 2.14: Monetary flow diagram of system with firms and households

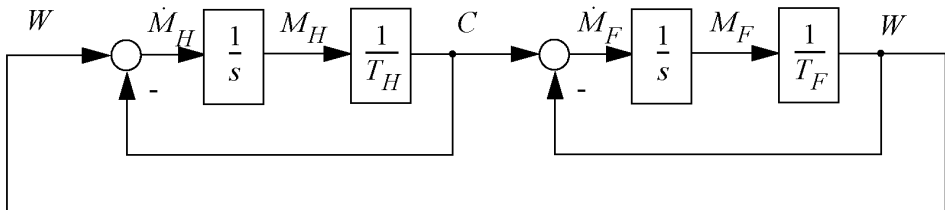


Figure 2.15: Elementary block diagram of firms/households system

Dependence on continuous time t is implied; we have $Y_d(t)$ and so on. Money stocks in the two sectors are M_F [\$] and M_H [\$]. As is clear from figure 2.14, we assume that there are no external sinks or sources of money. This assumption will be rescinded further below, among other things to discuss the multiplier.

The mathematical block diagram of this system is shown in figure 2.15. If we "reduce" (simplify) the two inner loop subdiagrams, using the rule in figure A.3, appendix A, we get figure 2.16.

There is one important loop lacking in this block diagram: the profit = investment loop depicted in figure 2.14. We should, however, note that figure 2.16 is entirely correct in the sense that in a system defined as consisting of firms and households, the input to the firm sector is consumption only,

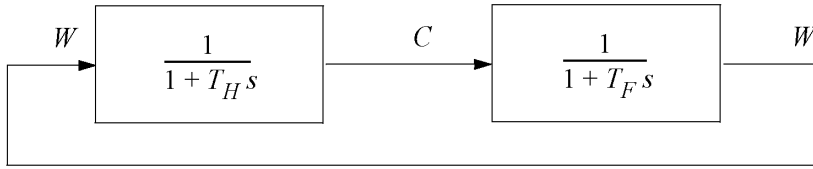


Figure 2.16: Equivalent block diagram of firms/households system

and the output is wages only – investment is a flow that is internal to the aggregate of firms¹⁴. So how do we extract profits, investment (and aggregate demand/output) from this block diagram representation? This is achieved by making the profit/investment flow loop external from the firms block in figure 2.16. We demand that the two firm blocks shown in figure 2.17 are equivalent in an input-output sense. This gives an equation to find

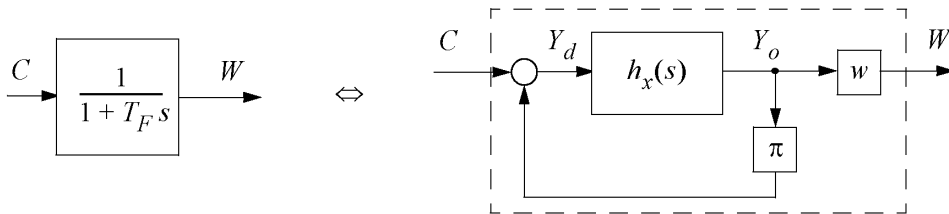


Figure 2.17: Extracting the profit flow loop

the unknown transfer function,

$$\frac{1}{1 + T_F s} = \frac{h_x w}{1 - \pi h_x} \quad (2.38)$$

Solving for h_x and using $\pi = 1 - w$:

$$h_x = \frac{1}{1 + w T_F s} \quad (2.39)$$

We observe that “extracting” the profit/investment loop leads to a reduced time lag $w T_F$ for the modified firm sector (this is the opposite operation of

¹⁴Here we have chosen to define the household aggregate to be of the worker type, not doing any investing. Later on we will define (or “extract” from the firm aggregate used here) “investing” or “capitalist” households in parallel with worker households – an alternative choice that is also useful.

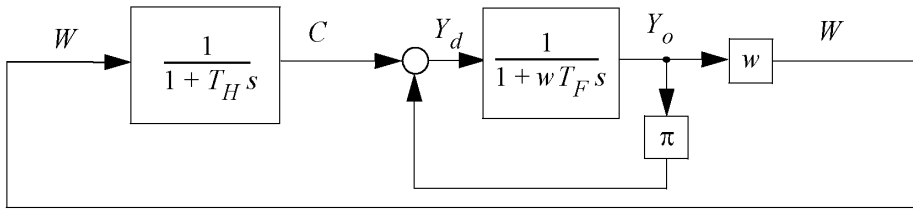


Figure 2.18: System with extracted profit loop

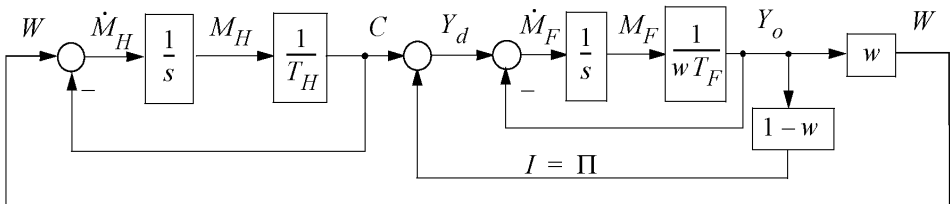


Figure 2.19: Elementary block diagram with extracted profit loop

incorporating a feedback loop, as given by (2.29)). Figure 2.16 may now be transformed into the equivalent block diagram shown in figure 2.18.

By making the profit/investment loop external to the firm sector, we are able to account for output Y_o in the block diagram. We may now expand figure 2.18 into an elementary block diagram corresponding to figure 2.15. The result is figure 2.19. (We have here substituted $1 - w$ for π). There are two integrators in this system. In other words we have a system with two states; household and firm money stock. This system is autonomous (it has no exogenous inputs), and its time path is therefore decided solely by the initial distribution of the money stock between the two sectors. We will now use this example to illustrate the use of Simulink, and then to find the equilibrium state of this system. The trajectory of the system is shown in figure 2.21. Initial values are assumed to be $M_{H0} = 1400$ and $M_{F0} = 800$ [\$]. System parameters chosen are time lags $T_H = 2$ and $T_F = 20$ [weeks], and wages' share of output is $w = 0.7$.

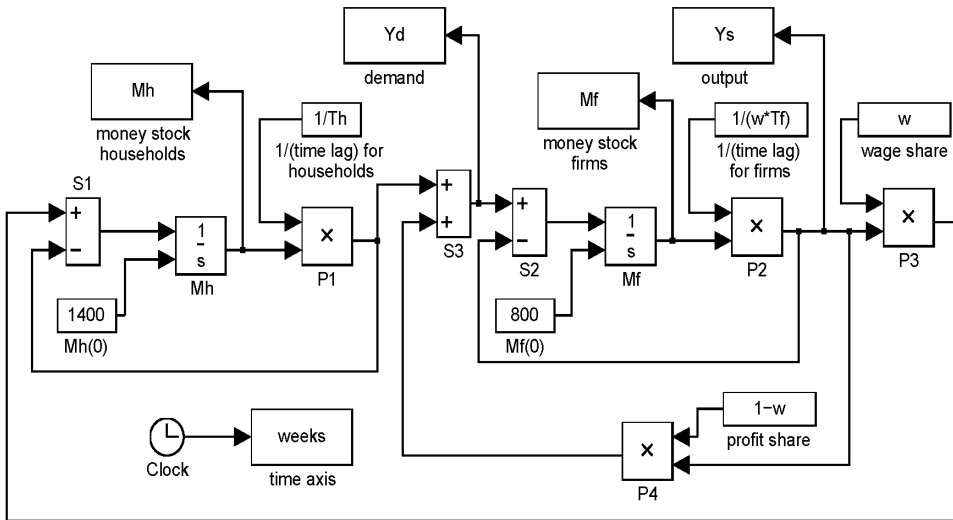


Figure 2.20: Simulink block diagram corresponding to figure 2.19

A Simulink block diagram¹⁵ corresponding to the one in figure 2.19 is shown in figure 2.20. This setup gives the responses shown in figure 2.21. We note how supply adjusts to demand in equilibrium. The graphs also indicate that in equilibrium, money stocks are proportional to the respective time lags in the two sectors. This is easy to see by considering figure 2.15: In equilibrium we must have $C = W$. This follows since

$$W = M_F/T_F \text{ and } C = M_H/T_H \quad (2.40)$$

Our main interest, however, is to focus not on equilibrium but dynamics. In this simple case we can find the algebraic solution for the trajectories, which is

$$M_F(t) = M_{F0}e^{-\alpha t} + \frac{M}{\alpha T_H}(1 - e^{-\alpha t}) \quad (2.41a)$$

$$M_H(t) = M_{H0}e^{-\alpha t} + \frac{M}{\alpha T_F}(1 - e^{-\alpha t}) \quad (2.41b)$$

¹⁵Note that summation points in this Simulink diagram are symbolised with rectangles with plus and minus signs, as opposed to circles used in the diagrams elsewhere in this paper and in most control engineering or signals and systems literature. The blocks with a cross ("X") imply that the output is the product of the inputs.

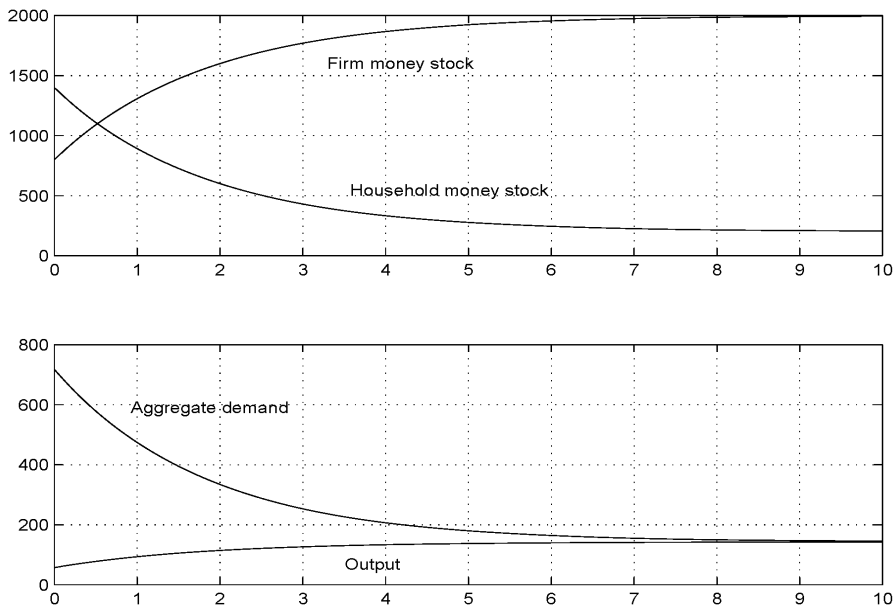


Figure 2.21: System path with no injection or drainage of money

where $\alpha = \frac{T_F + T_H}{T_F T_H}$. Total money stock, $M(t) = M_F(t) + M_H(t) = M_{F0} + M_{H0} = \text{constant}$, since there are no sources or sinks for money in this system. The system is linear and therefore amenable to algebraic solution. In a more realistic model with non-linearities, algebraic solutions are very difficult to find, if they exist at all. In such cases, numerical simulation is needed.

2.7.1 The economy simplified to a single first-order block

Consider the earlier discussed “tube” model in figure 2.6. We now want to replace the “tube” with a corresponding “vessel” (1st order time lag model), where – just as in figure 2.6 – an input demand flow $C + I$ results in an output flow which is recycled to the input. We will make use of the firm/households model in figure 2.18. The profit flow is now assumed to be paid out to stockholding owner households along with wages to workers, and households recycle the profit flow as investment to firms, in parallel with their consumption out of wages. This means that the profit flow is assumed to be lagged in the household sector with the same time constant T_H as the consumption flow. (This change is only done to make the presentation below simpler – a similar presentation could have been made with a model where

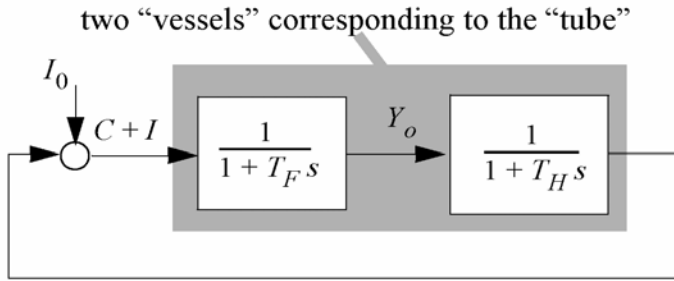


Figure 2.22: Vessel type model with output connected to input

investment out of profits are internal to the firm sector and not lagged, as in figure 2.16). We define a new $T_F := wT_F$ ¹⁶. The modified figure 2.18 can now be portrayed in a way that is structurally similar to figure 2.6. We have added an exogenous input I_0 , which might be flows from banks (loans), or injected via government spending. Note that the order of the firms and household blocks has been changed, so that the input to the first block (firms) is now $C + I$.

We may now as an approximation simplify the system to just one first-order time lag, using the series rule (2.20). This gives $\bar{T} = T_F + T_H$. Then we have a first-order block $\frac{1}{1+\bar{T}s}$ representing the shaded area in figure 2.22. The advantage of lumping firms and households together in this way is that we do not have to distinguish between lending, debt service, taxation and government spending for households, as opposed to firms. There is just one common injection (government spending, loans) and extraction (taxes, debt service) point, indicated by I_0 in figure 2.22. This simple model will be seen to be mostly sufficient for parts of this thesis.

.... and as an integrator

From the block diagram in figure 2.22 we find the transfer function from I_0 to Y_o , which is

$$\frac{Y_o}{I_0}(s) = \frac{\frac{1}{1+T_F s}}{1 - \frac{1}{1+T_F s} \cdot \frac{1}{1+T_H s}} = \frac{1}{\bar{T}} \frac{1 + T_H s}{1 + \frac{T_F T_H}{\bar{T}} s} \cdot \frac{1}{s} \quad (2.42)$$

¹⁶The notation $:=$ is used to signify a new interpretation of a symbol, related to the former use of the same symbol: “ T_F is from now on to be interpreted as w times the former T_F ”

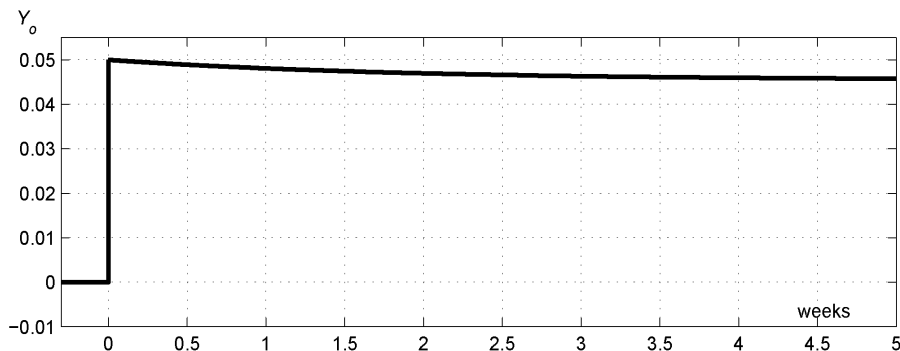


Figure 2.23: One-shot injection leads to sustained increase in Y_o

The transfer function contains a pure integrator $\frac{1}{s}$. A one-shot exogenous injection (an impulse) of money $I_0(t) = \delta(t)$ will give an output response as indicated in figure 2.23. Due to the integrator, this injected money leads to an output flow that jumps up and stays permanently higher than before the event. This is intuitively reasonable, since there are no leakages of money out of the circuit – any extra money injected must lead to sustained higher output. The increase¹⁷ in money stock is 1. To find the corresponding stationary increase in Y_o – which we will signify with $Y_o(\infty)$ – we may use the same argument as that leading to (2.40). Then $Y_o(\infty)$ is $\frac{1}{(T_F+T_H)} = 0.046$ (using the parameter values from the simulation of the model in 2.20). The not perfectly rectangular shape of the response is due to the transient process of the emerging flow from the firm sector having to fill up an initially empty household money stock, M_H . When this has been achieved, the system settles to a stationary state where $Y_o(\infty)$ is slightly lower, since some of the initially injected money now resides with households, and the rest with firms, with $Y_o(\infty) = M_F/T_F$.

We have $\bar{T} = T_F + T_H$. With $T_H \ll T_F$ (above we assumed 2 respectively 20 weeks). When output is fed back to the input, this results in the transfer function

$$\frac{Y_o(s)}{I_0(s)} = \frac{1 + T_H s}{\bar{T} s} \simeq \frac{1}{\bar{T} s} \quad (2.43)$$

¹⁷For simplicity, the simulation is done with money stock and other initial values = 0. This is acceptable since for a linear system dynamics are the same as if the system instead was started (more realistically) with initial values > 0 .

The Y_o impulse response is now a step (= perfect rectangular-shape) function. Its amplitude is the same as the stationary value in figure 2.23. With this ultimate simplification we see more clearly how the entire closed-loop economy observed from an outside money injection perspective is simply an integrator – which means a stock of money that is increased when additional money is injected. This is what remains when all loops are considered to be internal to the system (no loops are extracted).

If the outside injection $I_0(t)$ is not an impulse, but a persistent constant exogenous inflow of money [$\$/y$], money stock M and output $Y_o \approx M/\bar{T}$ will increase without bound. This is the case if the government runs a persistent deficit, which will be discussed in section 5.2.

2.7.2 Velocity and existence of money in the circuit

Augusto Graziani was a central researcher belonging to the *circuitist* school of economists, (Arestis and Sawyer; 2007). In several articles he discusses the effect of uncertainty, the velocity of money circulation and the seemingly strange question: why, actually, is there money in the circuit? He employs a simple circuit consisting of firms, households and banks:

Let us now assume a world free from uncertainty and populated by perfectly rational agents. In this world, any agent will go into debt only at the very moment in which he has to make a payment. Similarly, any agent who receives a money payment tries to spend it as soon as possible on goods or on securities. Both kinds of expenditure bring the money back to the firms, who immediately repay their debt to the bank. In a hypothetical world free from uncertainty and from frictions, the aforementioned steps would take place in an immediate succession with no time lag. This means that money is created, passed on from one agent to the next, and destroyed in the same instant. If this is the case, money is no longer an observable magnitude and the paradoxical result emerges of a monetary economy being defined as an economy in which money, in spite of its being by definition necessary for exchanges to take place, escapes any observation and any possible measurement. If all agents behaved as J.B. Say imagined, namely spending any amount of money as soon as received, the velocity of circulation would be infinite, money would be destroyed as soon as it was created and any attempt to measure the money stock in existence at any given moment

of time would invariably produce a zero result. As a paradoxical consequence, the image would emerge of a monetary economy (in the sense of an economy having ruled out barter and in which all payments are regulated in terms of money) in which money did not exist. (Graziani; 2003)

The problems pointed to by Graziani are simply and elegantly resolved by the use of time lag blocks in the circuit. The rules for aggregation of time lag blocks give us the relation between the amount of money in the circuit, the flow sizes, and money velocity. Here we use only firms and households, banks are not necessary for the argument. Consider figure 2.22. Imagine that after a while the exogenous injected flow I_0 is switched off. A constant amount of money M is now circulating. If we aggregate the household and firm block into one, the time lag is $\bar{T} \approx T_F + T_H$. Simplifying the system, as indicated in figure 2.24, we end up with the equation

$$Y_o(t) = \frac{M}{\bar{T}} = Mv, \text{ where } v \text{ is money velocity} \quad (2.44)$$

Money moves infinitely fast from the output and back to the input (in reality, from one account to the other), but stays within the aggregate block for a while, giving a finite \bar{T} , or $v = 1/\bar{T}$. With M constant, the upper system in figure 2.24 can be simplified down to just one integrator, generating the income flow $Y_o = M/\bar{T}$.

Further down Graziani writes:

For money to be an observable magnitude, it must be kept by single agents for a finite period of time, no matter how short, thus taking the form of a cash balance, be it notes or bank deposits. Since, as mentioned earlier, liquid balances are kept as a protection against uncertainty, this means that, for money to be an observable magnitude, the market must be operating under uncertainty. If we move in a hypothetical market free from uncertainty, liquid balances disappear, and with them the possibility of observing and measuring the money stock in existence.

Uncertainty is one reason for holding cash balances. But one does not need the assumption of uncertainty to justify that money velocity is not infinite – and finite money velocity is equivalent to agents holding buffers of money. It suffices that any agent, whether a firm or a household, needs to cope with the fact that outlays are not happening at the same time and

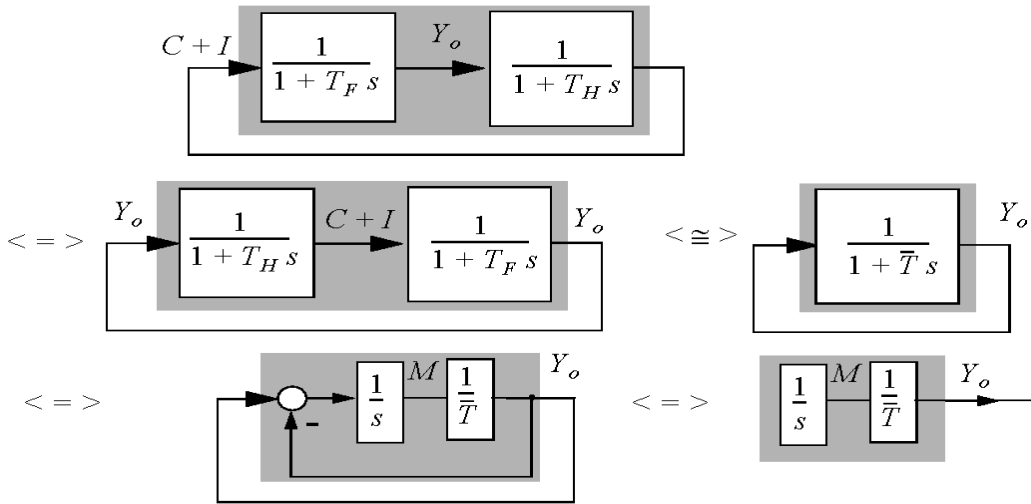


Figure 2.24: First-order blocks ensure finite money velocity and account for the existence of M .

with the same size as incoming amounts of money. For instance, a wage earner that receives money once a month, has to gradually portion out payments over the coming weeks, like portrayed by the impulse (input) and its response (output) in figure 2.13.

2.7.3 A time lag block with a small T is simply unity

Assume again the realistic case $T_H \ll T_F$, and consider the households block to the right in figure 2.22. The transfer function for this block is

$$h_H(s) = \frac{1}{1 + T_H s} = \frac{Y_{dH}}{Y_o}(s) \quad (2.45)$$

where we have introduced the symbol Y_{dH} for aggregate demand from households, i.e. the outflow from the household block. Eq. (2.45) corresponds to the differential equation

$$T_H \dot{Y}_{dH} = -Y_{dH} + Y_o \quad (2.46)$$

If we let $T_H \rightarrow 0$, the differential equation is reduced to $Y_{dH} = Y_o$, i.e. the outflow is now strictly equal to the inflow at any time, with no lag. We have instantaneous dynamics. This is reflected in the transfer function which

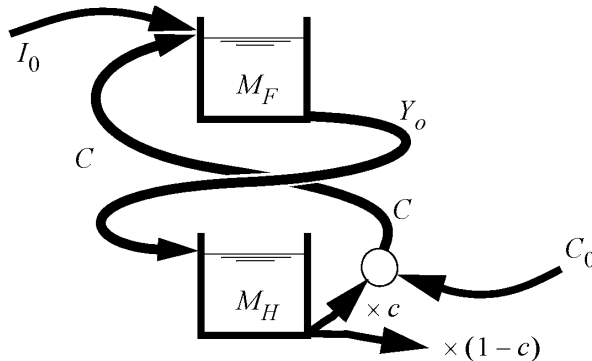


Figure 2.25: Money flow diagram with multiplier and exogenous inputs

becomes $h_H(s) = 1$. In block diagram symbolism we may then simply remove the households block and substitute it with a line connecting the former input and output, since such a line implicitly means multiplication by 1. It implies the simplifying but unrealistic assumption that households hold no money¹⁸.

2.7.4 The system with a multiplier

We will now discuss the multiplier. Historically, but also in today's economics textbooks, the multiplier was and is explained by partitioning the time axis into equidistant intervals (i.e. discrete time), and using a geometric series to derive how an exogenous injected flow of money leads to an increase in output that converges to a higher level. See for instance (Mankiw; 2008). We employ the common textbook model where all profits are paid to households together with wages, and where households consume a share c of their income and save the rest. See the monetary flow diagram in figure 2.25. The difference is that we do this in continuous time, and use the Laplace transform instead of a geometric series. We have two (exogenous) input money flows I_0 and C_0 , and a money flow is leaving the system as savings, through the savings coefficient $1 - c$. Using the vessel metaphor, money is persistently poured into the system at a constant rate, but drains out of the system also. The block diagram for the system, corresponding to

¹⁸This assumption is implicit in an attempted "dynamical extension" of the comparative statics IS-LM model which is critiqued in subsection 2.7.6 below.

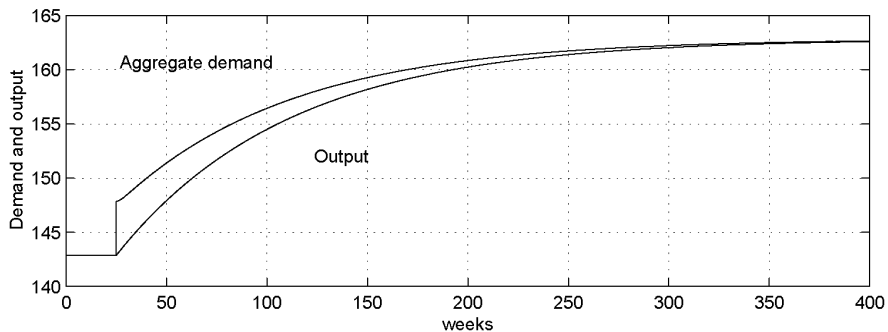


Figure 2.27: Time path with multiplier

nothing about the transient dynamic behaviour of the system (i.e. before equilibrium is reached).

We do not bring the algebraic solution for the system in in figure 2.26 here, but instead show the time path from a Simulink run, in figure 2.27. The system is initially in equilibrium with $C_0 + I_0 = 35.7$ when investment money flow is increased as a step function by $\Delta I_0 = 5$ at $t = 25$. The propensity to consume is assumed to be $c = 0.75$, i.e. we have a multiplier of 4. A 5 [\$/w] flow injection (as an exception we use *increase* in investment flow) results in output asymptotically increasing by 20 [\$/w].

The time of adjustment, estimated from the graphs, seems to be in the order of 80+ weeks. We may now use the rule (2.17) to check this by calculating the mean time lag of this response²⁰. To do this via the earlier derived rules, we first have to modify the transfer function (2.48) to be PUA:

$$(1 - c) \frac{Y_o}{I_0}(s) = \frac{1 + T_H s}{\frac{T_F T_H}{1-c} s^2 + \frac{(T_F + T_H)}{1-c} s + 1} \quad (2.50)$$

Even if the response for (2.50) has a smaller magnitude, it has the same shape and thus the same mean time lag. It is PUA, and we may use (2.17). We get

$$\bar{T} = a_1 - b_1 = \frac{(T_F + T_H)}{1 - c} - T_H = 86 \quad (2.51)$$

²⁰As opposed to this method, a drawback of the "classical" discrete-time method of using a geometric series to find the multiplier, is that the time constant of the process is lost. More generally: the loss – or abstracting from – actual and physical time is a disadvantage with time-discrete models.

The same result may be derived by using the series and feedback rules from subsection 2.6.4.

2.7.5 A Kalecki-inspired application

We will in this subsection use the developed approach for a model inspired by Michał Kalecki. One of his models is a simple circulatory system with capitalists and workers only. A famous quote credited to Kalecki by Joan Robinson and others (López G and Assous; 2010) is

Workers spend what they get, capitalist get what they spend.

We will present a model that expresses this. In addition, we will add an equation for the profit rate.

Readers are recommended to study the block diagram below, figure 2.28. All information needed is contained in that diagram. The model, however, will also be presented based on equations. We will see that it boils down to a three-state linear dynamical system.

We define the following variables and parameters; denominations are indicated in brackets:

T_F = time lag for the aggregate of non-financial firms [y]. There are no banks in this model.

T_W = lag for the aggregate of (non-saving) worker households [y].

$Y_o(t)$ = aggregate income to be shared between workers and capitalists [$\$/y$].

$Y_d(t)$ = aggregate demand to firms, for consumption and investment [$\$/y$].

$K(t)$ = capitalists' accumulated capital [$\$$].

d = depreciation rate on K [$1/y$].

r = profit rate [$1/y$].

π = share of aggregate income that capitalists receive []; $0 < \pi < 1$.
The workers' share is then $1 - \pi$.

σ = share of capitalists' profit flow that is invested, not used for consumption []; $0 < \sigma < 1$.

$I(t)$ = flow of investment [$\$/y$]

$W(t)$ = workers' wages flow [\$/y]; workers are assumed to use their entire wages for consumption = $C_W(t)$, with a lag²¹ T_W .

$\Pi(t)$ = profit flow to capitalists [\$/y].

$C_K(t)$ = capitalists' aggregate consumption flow [\$/y].

$C(t)$ = total consumption flow [\$/y], we have $C = C_W + C_K$.

In the following, dependence on time (t) is omitted for convenience. We start the presentation with the input/output dynamics of the two defined aggregates: firms and worker households. These dynamics may be explained via the firm aggregate which has time lag T_F – properties for the worker aggregates is similar, except for the time lag being T_W . We assume that all lags are of the first-order type, corresponding to a differential equation (using the aggregate of firms):

$$T_F \dot{Y}_o = -Y_o + Y_d \quad (2.52)$$

The money held at any time by the aggregate of firms, must satisfy

$$\dot{M}_F = -Y_o + Y_d, \quad (2.53)$$

so that

$$Y_o = \frac{M_F}{T_F} (= M_F v_F) \quad (2.54)$$

where v_F is firm money velocity [1/y] (but we will time lags in the following). With such input-output dynamics, a stepwise change in the input flow gives an output response that adjusts asymptotically to the input in the form of a stable exponential with a lag T_F . The worker household subsystem has similar properties. Equations are:

$$T_W \dot{C}_W = -C_W + W \quad (2.55)$$

with

$$C_W = \frac{M_W}{T_W} \quad (2.56)$$

The last differential equation is for capital accumulation and depreciation:

²¹We ignore the lag of capitalist households for simplicity. It might have been accounted for, but would not have made any difference for the arguments and conclusions.

$$\dot{K} = -dK + I \quad (2.57)$$

which is not part of the circuit but only a measure of success seen from the capitalists' position.

To complete the model, we need some additional (non-differential-) equations. The profit flowing to capitalists is:

$$\Pi = \pi Y_o \quad (2.58)$$

For workers' wages we have their share of output:

$$W = (1 - \pi) Y_o \quad (2.59)$$

For demand to firms we have:

$$Y_d = C + I \quad (2.60)$$

where

$$C = C_K + C_W = (1 - s)\Pi + C_W, \quad (2.61)$$

and

$$I = s\Pi, \quad (2.62)$$

This completes the set of equations describing the system.

The model until now described through a set of equations, is shown in figure 2.28 in the form of a block diagram:

This is a dynamic version of the Kalecki model. And we don't need to discuss how some equilibrium supposedly is achieved within a "period", since we use continuous time.

We now, however, wish to check whether a stationary system state is feasible, where the aggregate profit flow is positive and constant. We also want to examine which model parameters influence the profit rate, and how. Since the system is in equilibrium (= stationary), all derivatives are zero. From (2.53) we get the trivial result

$$Y_o = Y_d = Y \quad (2.63)$$

Capitalists invest the flow $I = \sigma\Pi = \sigma\pi Y$, cf. (2.62) and (2.58). They extract the flow $\Pi = \pi Y$. (As long as $\sigma < 1$, i.e. capitalists consume some of their profits, this corresponds to $m' > m$ in Marxian terms.)

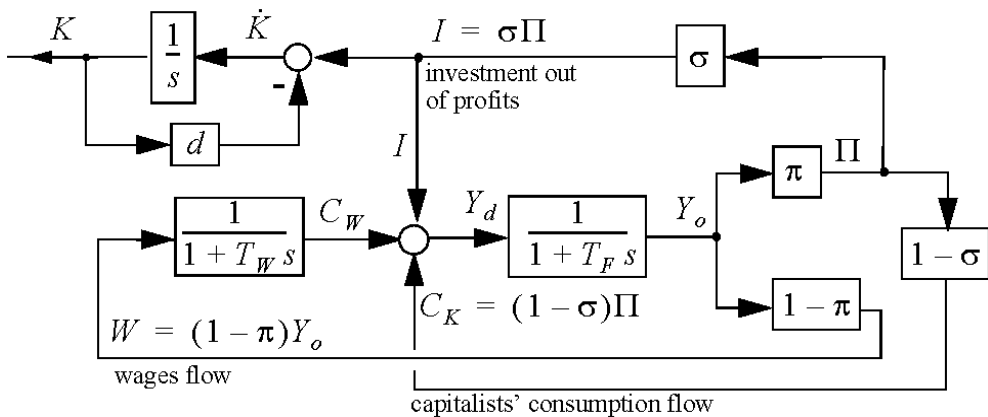


Figure 2.28: Block diagram for monetary circuit with capitalists and workers.

Turning now to the profit rate, and setting the left side of (2.57) to zero, we get

$$dK = I = \sigma\Pi = \sigma\pi Y \implies K = \frac{\sigma\pi Y}{d} \tag{2.64}$$

leading to the equilibrium profit rate

$$r = \frac{\Pi}{K} = \frac{\pi Y d}{\sigma\pi Y} = \frac{d}{\sigma} \tag{2.65}$$

Capitalists' profit rate in equilibrium is *not* dependent upon their profit share π of output. And the higher capitalists' consumption out of profits is (i.e. small σ), the higher the profit rate, giving support to the Kalecki quote. Capitalists also decide the profit rate in the sense that it is higher with a higher depreciation rate.

There are no banks in this model. But it is equivalent to a system with banks added, a special case where new loans plus banks's expenses are exactly equal to the debt service flow back to banks.

Two controversial conclusions emerge from this very simple model. The first is

1. *This economy can prevail without persistently increasing bank credit.* This is contested by many Post Keynesians, for instance Marc Lavoie, quote:

The debt of firms vis-à-vis banks must increase from period

to period, unless households decide to diminish their bank deposits.” Lavoie (1992), p. 156.

He also (ibid.) quotes colleague Paul Davidson:

... every repayment of the credit advanced by the authority must immediately be re-lent if activity is to be maintained.

In chapter 4 and later chapters we will see – using continuous time, not discrete "periods" – that it is feasible for an economy with bank lending to exist indefinitely even if new bank credit is not issued immediately when loans are repaid. It is also (theoretically) feasible for an economy to stably survive on a constant circulating money supply, as argued in this subsection. This presupposes falling prices with real growth (but this thesis does not go into price dynamics).

The second controversial conclusion from the above model is:

2. *The profit rate can be upheld regardless of long-term technological change.* This is contested by many Marxian economists, who subscribe to the theory of the persistently falling profit rate. This topic will not however, be discussed in this thesis.

2.7.6 An application: critique of a "dynamic" IS-LM model

We have up to this point tried to show how system-theoretic, block diagram-type tools are useful for macroeconomics, and to justify the first order time lag (“vessel”) model as a main component in such models. We will now use this and a well-known dynamic extension of the static IS-LM model to demonstrate that IS-LM as such is fundamentally flawed. IS-LM is to this day a central tool in macroeconomic discussion and decision-making, and is taught in macroeconomics undergraduate courses. Because of its central role there is also much critique of IS-LM. The brief analysis given below is also a critique, but more fundamental in the sense that it completely invalidates IS-LM. It is not based on arguments and considerations that may be more or less convincing depending on which economics camp one identifies with – but simply on a gross mathematical inconsistency, which if true cannot be contested.

We start with the static IS-LM equilibrium equations, where aggregate demand must equal output, $Y_d = Y_o = Y$; and demand for money, L must equal money stock, M . The interest rate is r :

$$Y = C(Y) + I(r) + G_0 \quad (2.66)$$

$$M = L(Y, r) \quad (2.67)$$

This is a simple IS-LM variant, with exogenous net government spending, and with investment being independent of output. But this simplified choice is sufficient for the arguments to be made. The model corresponds to the one given in Ferguson and Lim (1998), pp 2 - 3. The relations for consumption, investment and liquidity demand are assumed linear in output and/or interest. Then we have

$$Y = C(Y) + I(r) + G_0 = C_0 + cY + I_0 - br + G_0 \quad (2.68)$$

$$M = L(Y, r) = kY - hr \quad (2.69)$$

Here c, b, k, h are constant parameters > 0 .

We remind ourselves at this stage that this "comparative statics" model has as its premise that it is a simplified representation; it is an "equilibrium-in-each-period version" of something that in reality is continuously varying dynamic system. John Hicks, the inventor of the IS-LM model, writes:

The IS-LM diagram, which is widely, but not universally, accepted as a convenient synopsis of Keynesian theory, is a thing for which I cannot deny that I have some responsibility. It first saw the light in a paper of my own, "Mr. Keynes and the Classics" [1937], but it was actually written for a meeting of the Econometric Society in Oxford in September 1936, just eight months after the publication of *The General Theory*

....

... one can hardly get a plausible rule while confining attention to what happens within a single period. So it would seem that the proper place for such a proceeding is in sequential models, composed of a succession of periods, in each of which the relevant parameters have to be determined; there is then room for linkages between the periods, and so for lags. I have myself made some attempts at the construction of such models. I think they have their uses, but they are not much like IS-LM.

If one is to make sense of the IS-LM model, while paying proper attention to time, one must, I think, insist on two things: (1) that the period in question is a relatively long period, a "year" rather than a "week"; and (2) that, because the behavior of the

economy over that "year" is to be determined by propensities, and suchlike data, it must be assumed to be, in an appropriate sense, *in equilibrium* (Hicks; 1981).

Recognising the weaknesses of "statics within a period", Ferguson and Lim attempt to construct a "dynamic extension" of this model:

$$\dot{Y}_o = \alpha(Y_d - Y_o) = \alpha(C_0 + cY_o + I_0 - br + G_0 - Y_o) \quad (2.70)$$

$$\dot{r} = \beta(L - M) = \beta(kY_o - hr - M) \quad (2.71)$$

α, β are constant parameters > 0 . Verbally, these two differential equations state that the rate of change of output is proportional to the difference between aggregate demand and output, and that the rate of change of the interest rate is proportional to the difference between demand for liquidity²² and money stock. The denominator for the stock M is still [\$], while Y_d, Y_o, C_0, I_0, G_0 now get the denomination [\$/y] and become *flows* – in contrast to their denomination in the comparative statics model, which is [\$/]. We will now argue that this model is fundamentally flawed, in spite of being dynamic.

We represent equation (2.70) by a block diagram, see figure 2.29²³.

For the block diagram corresponding to the money market equation (2.71), see figure 2.30.

Before combining these two diagrams to one representing the whole system, we reformulate equation (2.70). It may be written as

$$\dot{Y}_o/\alpha = (Y_d - Y_o), \text{ which obviously must be } = \dot{M}_F, \quad (2.72)$$

since $Y_d - Y_o$ is the net money flow into the firm sector. By this we have incorporated firm money stock in the model. Equation (2.72) explains the slightly reformulated but equivalent "firm" substructure in figure 2.31 below, which – except for this modification – is a result of a straightforward

²²One could reasonably argue that the transaction demand for money in (2.71) should instead be kY_d , but the choice is to follow Ferguson and Lim. And this choice does not have any impact on the argument to be made.

²³Note that this dynamic model implies that the household sector has instantaneous dynamics, signified by the block with unity gain. Comparing with figure 2.22, this corresponds to the time lag in the household sector being set to zero. This assumption may be acceptable, since the time lag of the firm sector is much larger in comparison, due to a high amount of between-firm transactions, which is discussed later in section 2.8. One should however, be aware that this assumption implies that money stock in the household sector is zero: there is no buffer there, only a through flow.

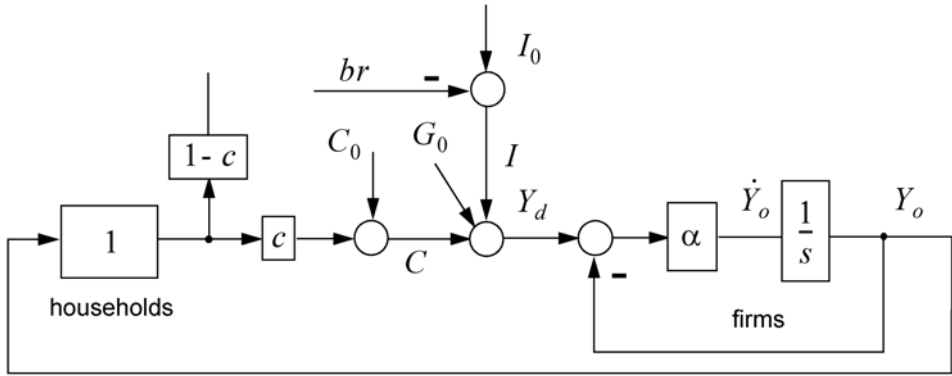


Figure 2.29:

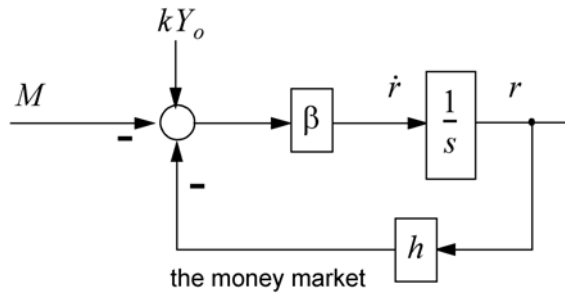


Figure 2.30:

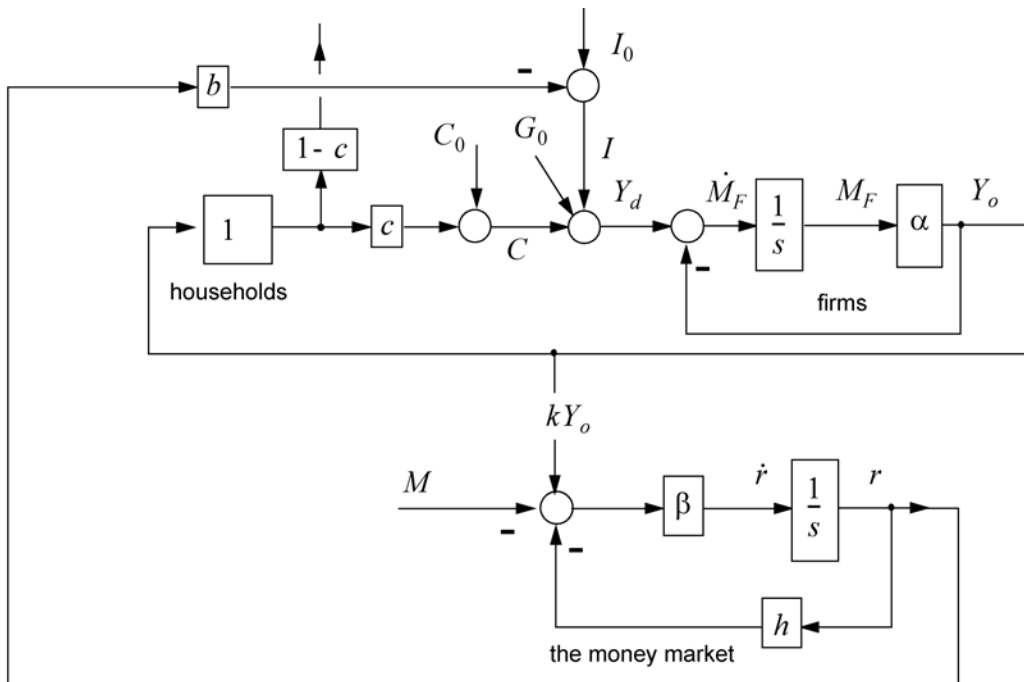


Figure 2.31:

connection of the two sub-diagrams for the real economy and the money market²⁴.

By this modification we have accounted for the dynamics of the firm sector money stock, which in fact *must be identical to the entire money stock of the economy*, since households in this model are implicitly assumed to have no money stock, and the financial sector only appears indirectly via exogenous flows.

By now the inconsistency of the IS-LM model may be clearly observed: While money stock in reality is endogenous (M_F) and a system state, it is at the same time assumed to be an exogenous (input) variable (M). What makes this inconsistency go unnoticed, is that the actual presence of money stock M_F within the the Y_d to Y_o dynamics, *disappears in the (comparative) statics framework*.

²⁴The modification (2.72) may alternatively be explained by exploiting a rule for block diagram manipulation: Interchanging the sequence of blocks on a path (in this case the two blocks containing α and $\frac{1}{s}$) does not change the transfer function along that path.

set its target at 25 basis points, or raise it to 150 basis points. Finally, the control sense and the theoretical sense are related but not identical. Let us say a country has a fixed exchange rate and uses the interest rate policy to hit the peg. We can say the interest rate is exogenously controlled (set by the Central Bank) but it is not theoretically exogenous because the overriding policy is to peg the exchange rate. In the theoretical sense, the Central Bank's concern is to hit the exchange rate target so that it has surrendered control of the interest rate (it uses the interest rate as a tool to hit the targeted exchange rate). On the other hand, let us say that the Central Bank targets full employment and uses the interest rate to achieve that target. Again we would say the interest rate is exogenous in the control sense, but not in the theoretical sense because it is used to target full employment, (Wray; 2012, section 3.5)

Wray's qualifications that the interest rate is not freely set by the Central Bank, but is often used to accommodate needs related to the exchange rate or employment, does not invalidate the fact that when the Central Bank sets its target rate based on some policy considerations, this is decisive for the other rates in the financial sector. These rates are not decided by "demand" for liquidity, as assumed in IS-LM.

2.8 Aggregating agents into a sector

We now choose the first-order time lag model as a description of the behaviour of an individual "micro" agent; a single household or firm. The generic symbols $F_i(t)$ and $F_o(t)$ signify in- and outflow of money. Money stock $M(t)$ may – as stated earlier – be interpreted as the agent's necessary liquid buffer to handle discrepancies between in- and outgoing money flows.

The agent is, just as in the aggregate case, assumed to react to a monetary step function income flow with a time-dispersed exponential spending response asymptotically approaching the incoming flow level, as already depicted in figure 2.2. We repeat the equations for the first-order time lag model, but use τ to distinguish it from T_a , which shall be reserved for the lag of the aggregate of agents:

$$\dot{M}(t) = -F_o(t) + F_i(t) \quad (2.73a)$$

$$F_o(t) = \frac{M(t)}{\tau} = M(t)v \quad (2.73b)$$

The agent's transaction frequency (money velocity) is $v = 1/\tau$, which is also a measure of confidence, a topic which we will return to later. The parameter τ (or the inverse, v) is our first behavioural assumption for the generic agent. One may let τ be influenced by other system variables, for instance it could increase sharply due to falling confidence/optimism in a recession/depression. Such modifications will make a model containing such agents nonlinear. But for the time being we will stick to the assumption of a constant τ for all agents, which gives a linear system. The τ 's are also assumed identical.

2.8.1 An aggregation theorem

We assume that our aggregate consists of a large number of identical individual agents as described above. An aggregate of agents (a sector) may typically represent all firms, as in Phillips' macroeconomic model. The individual agents that constitute a given sector will of course have different "sizes" in the sense that money stock and flow magnitudes will vary widely between them. But we assume that (2.73) holds for all agents in a given aggregate, i.e. that the outflow from an agent is proportional to the agent's money stock, by a common constant velocity $v = 1/\tau$ (the assumption of a common v will be relaxed later on). Thus all agents in the sector is represented by the transfer function

$$h(s) = \frac{1}{1 + \tau s} \quad (2.74)$$

We further assume that any (in an average sense) individual agent's outgoing money flow is divided into a share ρ (out of the sector) and $1 - \rho$ (to other agents within the sector), where $0 < \rho \leq 1$. We will call ρ an *outflow coefficient*, or *outside spending coefficient*. See the flow diagram in figure 2.33. The shaded arrows indicate a network of interactions, where any individual agent in principle interacts with any other agent. Our interest is focused on two aspects, input-output characteristics of the aggregate, and the dynamics of aggregate money stock.

Under the above assumptions, the transfer function for the sector turns out to be surprisingly simple:

Theorem 2.1 *Given a network of an infinite number of identical units which are all first order transfer functions (2.74), and which have identical outflow coefficients ρ , and where the outflow share for each transfer function that goes to all other agents, sum to $1 - \rho$. Then the transfer function for the*

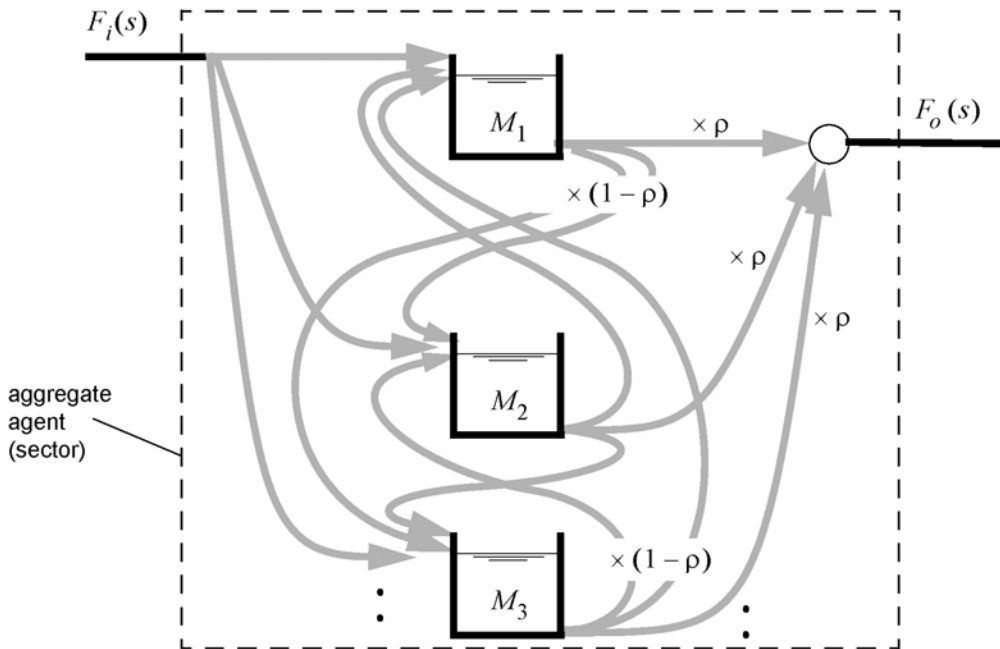


Figure 2.33: A flow network of “vessel” agents

network, regardless of how the input to the network is partitioned between agents in the network, is

$$h_a(s) = \frac{F_o}{F_i}(s) = \frac{1}{1 + T_a s}, \text{ where } T_a = \frac{\tau}{\rho} \quad (2.75)$$

The proof is given in (Andresen; 1998), but will be repeated here. Before proceeding, some comments to indicate that this result is intuitively satisfying. Let us first consider a type of sector where the population of agents have a low volume of monetary transactions between them, even if the number of agents may be large: A case in point is the aggregate of all households. In this case ρ is close to unity. Referring to figure 2.33, this means that the agents simply act “in parallel”, with negligible flows between them. Money received by a specific agent will emerge from the the agent and also the household sector, without having to “detour” via other household agents first. People use most of their income for purchasing goods and services from firms, not paying it to other households. Thus one should expect

the aggregate of households to have the same fast response as an individual agent. This fits with (2.75), since $T_a = \tau$ when $\rho = 1$.

The other extreme is when the “aggregate agent” is such that agents mostly do their transactions with other agents within the aggregate. This case fits well with what financial sectors have developed into for the last decades. An outside agent who injected money into such an aggregate, would – if she had the means to trace that amount of money – observe that it would take a very long time before the last residue of the injected amount emerged from the aggregate. This case corresponds to ρ being close to zero. It is consistent with (2.75), where a small ρ means a large lag T_a , giving just the type of low-amplitude, drawn-out response that seems reasonable. We call this the *detour effect* from now on.

We will now prove Theorem 2.1.

Proof. In deriving the transfer function for the aggregate agent, we may assume that the outside incoming monetary flow arrives at one agent only, because of the symmetry between the agents, and because of the superposition principle that applies to a linear system: If the incoming flow was instead distributed between several agents, the resulting response would be the sum of responses to each component of the incoming flow, transmitted through identical transfer functions, which would then sum up to the same result we get when the incoming flow is assumed to arrive at a single agent only.

Consider the structure in figure 2.34. This block diagram accounts for the way an incoming monetary flow branches through the aggregate of agents. As already argued we may assume that the entire inflow enters at one single agent without loss of generality – in figure 2.34 chosen as the top agent. This results in an outflow from that agent which is partitioned into a share ρ leaving the aggregate, and a share $1 - \rho$ to another identical agent within the aggregate. The latter flow again results in a flow that is partitioned into a share leaving the aggregate, and a share to another agent within the aggregate, and so on. Feedback loops are indirectly accounted for by the structure in figure 2.34, since the effect of any feedback loop may be equivalently represented by an infinite succession of series and parallel connections through identical transfer functions. For the transfer function for the aggregate we have

$$h_a(s) = \frac{F_o}{F_i}(s) \quad (2.76)$$

as indicated in figure 2.34 by the light shaded area. If we now remove the upper single agent from the aggregate, and assume that the remaining number of agents is so large that this does not significantly affect the dynamics of the

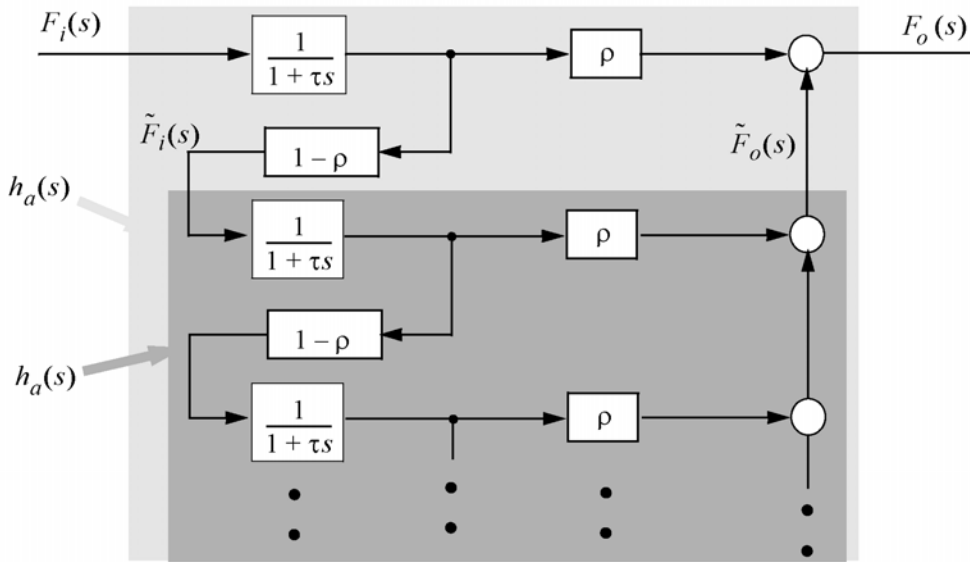


Figure 2.34: Block diagram of an input-output equivalent network that is without feedback loops

aggregate, then $h_a(s)$ may also be found as indicated by the dark shaded area, so that

$$h_a(s) = \frac{\tilde{F}_o(s)}{\tilde{F}_i(s)} \quad (2.77)$$

Employing rules for manipulating block diagrams where blocks are in parallel and in series, we get

$$h_a(s) = \frac{\rho}{1 + \tau s} + \frac{1 - \rho}{1 + \tau s} h_a(s) \quad (2.78)$$

Solving for $h_a(s)$, we get (2.75). ■

For use later on we will give an alternative proof:

Proof. Consider the block diagram in figure 2.35. As already defined, $h(s) = 1/(1 + \tau s)$. Again we assume that all transfer functions in the network are identical so that the incoming money flow to the network may be considered to arrive at one agent only. A unit of money arriving at the input will first be lagged through the upper transfer function $h(s)$. Then a share ρ leaves the system, and the remaining share $1 - \rho$ is lagged through an identical transfer function (the lower one). The output from this transfer

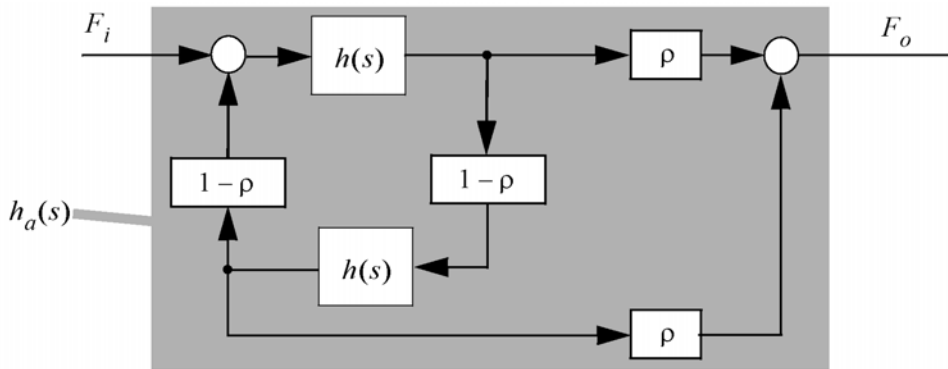


Figure 2.35: An input-output equivalent network with one feedback loop

function is split in a similar way, and the remainder is input to the upper transfer function, lagged and split, and so on. Thus the feedback structure must give the same aggregate input/output-dynamics as the structure in figure 2.34. Using block diagram manipulation and reduction rules, the transfer function from input to output in figure 2.35 must then be

$$\frac{F_o}{F_i}(s) = h_a(s) = \frac{\rho h}{1 - (1 - \rho)^2 h^2} \{1 + (1 - \rho)h\} = \frac{\rho h}{1 - (1 - \rho)h} \quad (2.79)$$

Substituting $h = 1/(1 + \tau s)$ and simplifying, gives (2.75). ■

From the intermediate result (2.78) follows

$$h_a(s) = \frac{\frac{\rho}{1 + \tau s}}{1 - \frac{1 - \rho}{1 + \tau s}} \quad (2.80)$$

This corresponds to the block diagram shown in figure 2.36. It may be formulated as an “equivalent simplified structure” corollary:

Corollary 2.2 *Given a network as defined in Theorem 2.1. The associated input-output dynamics are equivalent to that of a system with one agent $\frac{1}{1 + \tau s}$ where an outflow share $1 - \rho$ is fed back to the input and a share ρ flows out of the system.*

This is intuitively satisfying, since we may look upon this structure as arranging all identical agents in parallel and bundling all their outflows that go to other agents into one common feedback flow, while the remaining flows leave the system as another aggregated bundle.

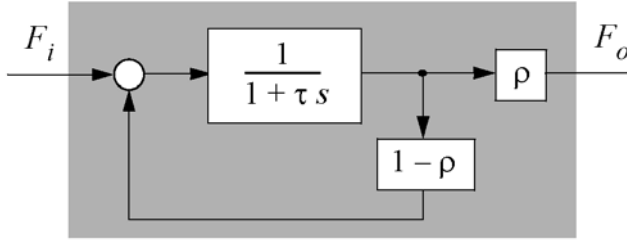


Figure 2.36: Equivalent structure corresponding to network of agents

From Theorem 2.1 follows a “look-inside corollary”:

Corollary 2.3 *Given a network as defined in Theorem 2.1, and an output flow $F_o(t)$ from the network. Then the sum of internal flows within the network is*

$$F_w(t) = \frac{1 - \rho}{\rho} F_o(t) = \frac{T_a - \tau}{\tau} \quad (2.81)$$

Related to this, we have a “loop extraction corollary” – which was already implied in (2.39):

Denote the time lag of a given aggregate T . Assume that one wants to “extract” a bundle of internal flows summing to F_{wi} and then add it to the originally defined outflow F_o , so that the redefined outflow is $\hat{F}_o = F_o + F_{wi}$. Assume that the coefficient $w < 1$ is given so that $\frac{F_o}{F_{wi}} = \frac{w}{1-w}$. Then the transfer function $\frac{\hat{F}_o}{F_i}(s)$ becomes

$$\frac{\hat{F}_o}{F_i}(s) = \frac{1}{1 + wTs} \quad (2.82)$$

This is what we did in connection with figure 2.17.

The opposite is “loop inclusion”:

Corollary 2.4 *Call the time lag of a given aggregate T . Assume that one wants to include or “hide” a share $1 - w$ of the outflow F_o inside the aggregate, so that the remaining outflow is $\hat{F}_o = wF_o$. Then the transfer function $\frac{\hat{F}_o}{F_i}(s)$ becomes*

$$\frac{\hat{F}_o}{F_i}(s) = \frac{1}{1 + \frac{T}{w}s} \quad (2.83)$$

Another result is a “liquidity preference corollary”:

Corollary 2.5 *Assume that liquidity preference is increased, which in our model is expressed by a larger τ . This means that both aggregate output flow and the aggregate of internal flows will be lower, since*

$$F_o(t) = \rho \frac{M(t)}{\tau} \text{ and } F_w(t) = (1 - \rho) \frac{M(t)}{\tau} \quad (2.84)$$

Such a mechanism is at work during serious economic downturns. Decreased flows may lead to a more pessimistic mood leading to further increase in τ , which – following (2.84) – means further decreased flows, and so on. We may get an unstable process towards a depression. This is explored in chapter 6.

Closely related to corollary 2.5 is this, on "internal flow depletion":

Corollary 2.6 *Assume that there is an imposed lower bound $F_{oL}(t)$ on the outflow $F_o(t)$ with $\dot{F}_{oL}(t) \geq 0$, and that the inflow $F_i(t)$ is not increasing at the same rate. Then the sum of internal flows must decrease relative to $F_o(t)$ ($\iff \rho$ must increase) if individual agents' time lags τ are not decreased to compensate.*

Such imposed or “non-discretionary” flows for a sector (or an agent) may typically be taxes or debt service. But it is perhaps more correct to let non-discretionary flows be accounted for by subtracting them at the *input* of the unit in question. The unit simply has to abide with them, it is not a choice of the unit. We will choose this alternative in the rest of this thesis.

2.8.2 Allowing for differing agent transfer functions

We will now relax some restrictions: all agents are still PUA and have the same time lag and same outflow coefficient ρ . Except for that, agent transfer functions and the corresponding impulse response shapes may now be completely arbitrary and different. Consider figure 2.34, but with such more general transfer functions in the blocks. We apply the series connection rule (2.20) and the parallel connection rule (2.28) to a structure corresponding to the intermediate result, equation (2.78). This gives

$$T_a = \rho\tau + (1 - \rho)(\tau + T_a) \quad (2.85)$$

Solving for T_a gives $T_a = \tau/\rho$, as expected.

2.8.3 Example: Lag of a sector with consumption goods, intermediate and investment goods

We consider a model where the sector of firms is partitioned into two categories (two subsectors): Consumption goods firms ("C") on one hand, and firms producing intermediate or investment goods ("I&I") on the other. Figure 2.37 shows the flow structure. For this example, we define both investment

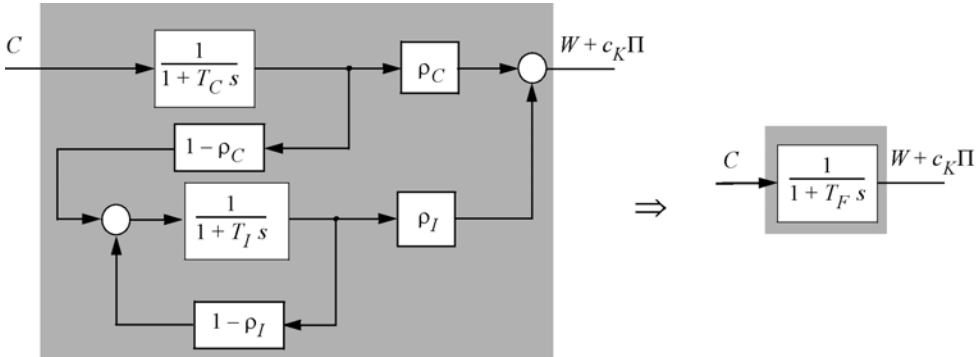


Figure 2.37: Flows in a firms sector with a consumption goods subsector and an intermediate/investment goods subsector.

and intermediate goods as flowing internally in our firm sector, so that the incoming flow to the sector as a whole is consumption only (from workers who are assumed not to save/invest, plus capitalists' consumption)²⁵. Out-flow coefficients are ρ_C and ρ_I as indicated in the figure. C firms buy I&I goods indicated by the flow with a factor $1 - \rho_C$, and the I&I subsector purchases from itself indicated by the flow with the factor $1 - \rho_I$. The lags of the two subsectors are T_C and T_I . The outgoing flow is wages plus profits to be used for consumption. We now seek the lag T_F of the firm sector as a whole, which we will expect to be greater the smaller ρ_C and/or ρ_I are. We apply two of the rules developed in section 2.6. First we find the effect on the lag of the I&I sector of the sector's purchase flow from itself through the factor $1 - \rho_I$. Corollary 2.4 (loop inclusion) in section 2.8 gives the time lag T_I/ρ_I . Now we are left with two parallel branches which taken together are serially connected to the subsector with time lag T_C . The upper parallel branch with coefficient ρ_C has zero lag. Using rule (2.28) the lag of the parallel branches taken together is then $(1 - \rho_C) \cdot T_I/\rho_I + \rho_C \cdot 0 = T_I \cdot (1 - \rho_C)/\rho_I$.

²⁵The implicit assumption here is that all investment is made out of retained earnings, nothing from bank loans. For the point to be made here, this is acceptable.

The next and last step is to include the subsector with time lag T_C . We use the series connection rule (2.20) which gives

$$T_F = T_C + \frac{1 - \rho_C}{\rho_I} T_I \quad (2.86)$$

The "detour effect" is more pronounced the smaller the outflow coefficients are (i.e. the closer to unity the feedback coefficients are). This means a larger time lag for the firms sector as a whole. This is confirmed by (2.86).

2.9 Aggregation check through Monte Carlo simulations

2.9.1 Network of units with rational transfer functions

We will now examine the dynamics of the aggregate of a large number of agents, chosen such that they differ in several ways. Results cannot be reached algebraically and have to be worked out through numerical simulation. We assume that a PUA second order rational transfer function with one zero and real poles is sufficient for furnishing the necessary variability in individual agent properties. We express the agent PUA transfer function in the form

$$h(s) = \frac{\alpha_1 + \alpha_1 T_z s}{\alpha_1 + \alpha_2 s + s^2} = \frac{\alpha_1(1 + T_z s)}{(s + 1/T_{p1})(s + 1/T_{p2})} \quad (2.87)$$

The point of the zero $-1/T_z$ is to account for some agents spending a certain share of incoming money immediately, when received. For a large T_z , the agent's spending reaction will start with a fairly strong initial spike followed by a correspondingly small exponential tail. Figure 2.38 shows a collection of spending impulse responses for 10 agents, where all have a transfer function of the type (2.87). They are all PUA, and they have the same time lag, here $\tau = 1$. The same set of responses is shown in both windows, but with different scaling. In the lower left window we note the large initial amplitude of some responses; these are the cases with a fairly large T_z . The first-order time lag response is shown with a thick dotted line for comparison, and its tangent (and τ) is also indicated.

Each response corresponds to a specific parameter set T_z, T_{p1}, T_{p2} where $-1/T_{p1}$ and $-1/T_{p2}$ are the poles of (2.87). Each set is generated by uniform probability density functions, through the following procedure: First a T_z is generated in the range $0 < T_z < 0.5\tau$ (the factor 0.5τ is fairly ad hoc).

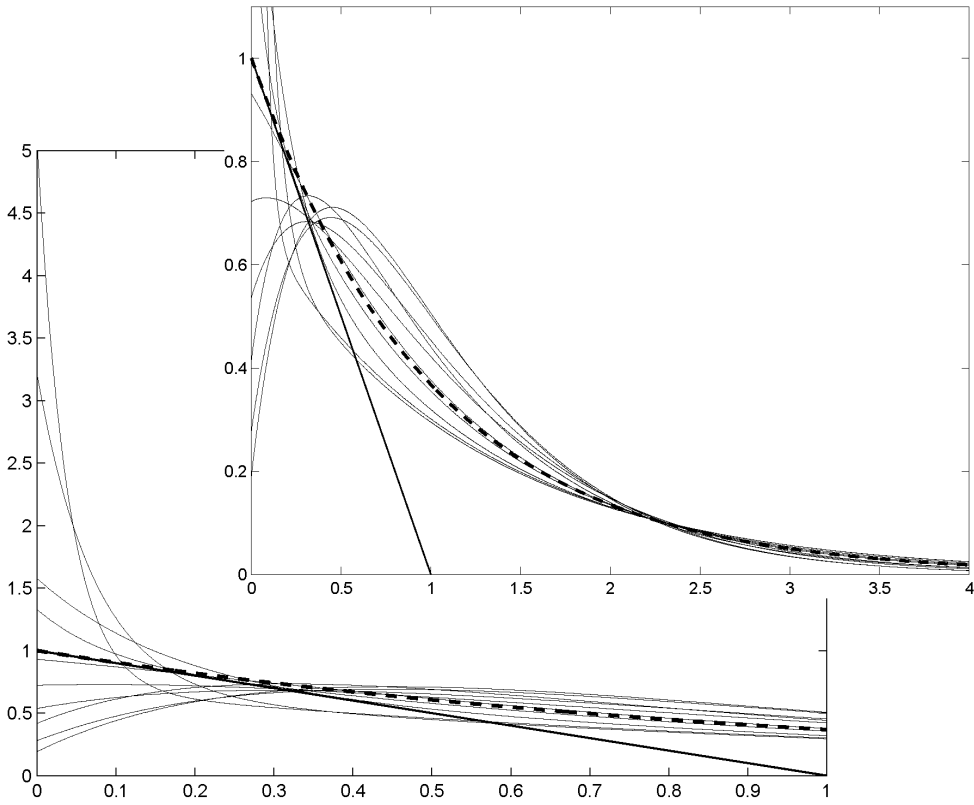


Figure 2.38: Ten impulse responses of (2.87), all with $\tau = 1$

According to (2.17) and (2.20), the sum of the denominator time constants must be $T_z + \tau$. The next step is generating the pair T_{p1}, T_{p2} via drawings from a uniform p.d.f., but scaled afterwards such that this condition is satisfied. We then have $\alpha_1 = \frac{1}{T_{p1}T_{p2}}$ and $\alpha_2 = \frac{1}{T_{p1}} + \frac{1}{T_{p2}}$. These α 's are used in the matrix A below. This procedure ensures that any response generated will be PUA.

We observe in figure 2.38 that the first order time lag is not a good approximation to most of the responses that are generated. Our conjecture however, is that if we interconnect a large number of agents with differing responses, we will observe that the first order time lag as an approximation for a sector improves with

1. the number of interconnected agents that constitute the sector.
2. the degree of interconnectedness between these agents.

To explore this, a state space model is defined. It consists of subsystems of the type (2.87), interconnected such that the total system is still PUA. The state space model is realized as a controllable canonical form (see for instance (Belang er; 1995), pp. 100-104). Let

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \quad (2.88a)$$

$$y = \mathbf{c}^T \mathbf{x} \quad (2.88b)$$

The matrix $A =$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ -\alpha_{11} & -\alpha_{12} & r_{21}\alpha_{21} & r_{21}\alpha_{21}T_{z2} & r_{31}\alpha_{31} & r_{31}\alpha_{31}T_{z3} & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ r_{12}\alpha_{11} & r_{12}\alpha_{11}T_{z1} & -\alpha_{21} & -\alpha_{22} & r_{32}\alpha_{31} & r_{32}\alpha_{31}T_{z3} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (2.89)$$

Each subsystem has transfer function (2.87). The subsystems (agents) are indexed $i = 1, \dots, n$. With n subsystems the state space model has order $2n$. The coefficients α_{i1} , α_{i2} and T_{zi} correspond to those in (2.87). Note that indices in (2.89) are not matrix element indices since they pertain to the subsystems. Each coefficient r_{ij} accounts for the flow from subsystem i to subsystem j . For each subsystem i there is an outflow coefficient ρ_i . Then r_{ij} must satisfy

$$\sum_{j=1, j \neq i}^n r_{ij} = 1 - \rho_i \quad (2.90)$$

The column vector \mathbf{b} has $2n$ elements and is $\mathbf{b} = [0 \ b_1 \ 0 \ b_2 \ \dots \ 0 \ b_n]$. It must satisfy

$$\sum_{i=1}^n b_i = 1 \quad (2.91)$$

Conditions (2.90) and (2.91) are necessary to achieve unit area impulse response for the total system, i.e. such that no money is created or destroyed within it.

The row vector \mathbf{c}^T has $2n$ elements and is

$$\mathbf{c}^T = [\rho_1 \alpha_{11} \ \rho_1 \alpha_{11} T_{z1} \ \dots \ \rho_n \alpha_{n1} \ \rho_n \alpha_{n1} T_{zn}], \text{ all } \rho_i < 1 \quad (2.92)$$

Now to the procedure for assigning values to the above parameters: for each Monte Carlo run, a complete new set is generated. All probability distributions employed are uniform. The procedure is executed for each subsystem i : First, a lag τ_i is drawn. Based on this, the parameters α_{i1}, α_{i2} and T_{zi} are generated as already described. Then an outflow coefficient $\rho_i < 1$ is drawn. Next, coefficients r_{ij} are also drawn, but scaled afterwards such that condition (2.90) is satisfied. Now we also have the next two elements in \mathbf{c}^T . After repeating this for all subsystems, all elements in \mathbf{b} are drawn and then scaled such that (2.91) is satisfied.

One should expect widely differing responses, since all parameters are allowed to vary quite independently, and the distributions employed are chosen to have a fairly wide range. This makes the test of our conjecture more severe. The ranges chosen are

$$0.1 < \tau_i < 1.9, \ 0 < T_{zi} < 0.5\tau_i, \ 0.1 < \rho_i < 0.7 \quad (2.93)$$

The mean of τ_i is then 1.0 and the mean of ρ_i is 0.4. Based on this, the approximate first order time lag impulse response for the total system is predicted to be

$$h(t) = 0.4e^{-0.4t} \quad (2.94)$$

We start simulations with a system of only 10 agents. The lower left window in figure 2.39 shows the responses for each individual agent. We note that the responses, as opposed to those in figure 9, are much more dispersed now since also the time lags differ between them (in the range 0.1 to 1.9). The mean first order response and its tangent is also indicated. The upper right window shows the impulse response of the ‘‘sector’’ consisting of these ten agents, together with the first order time lag response. We observe that this proposed approximation is not too bad, as predicted.

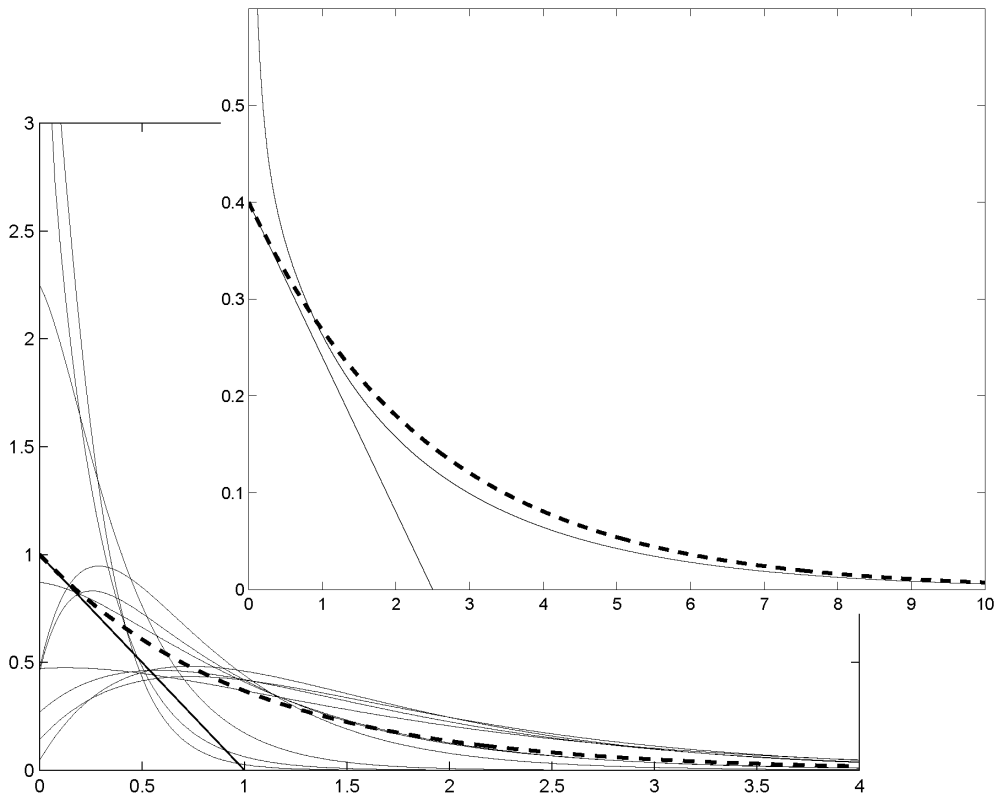


Figure 2.39: Ten individual responses (LL), and the corresponding aggregate response (UR)

A population of as little as ten agents in a sector is quite unrealistic. We therefore do the same with a 150-agent system, which means a 300×300 system matrix. In figure 2.40 are given 10 impulse responses for a 150-agent system, in the lower left window. For comparison, 10 responses for a 10-agent system are given in the upper right window. We observe that the first order approximation is better for the sector with the higher agent population.

Now to the impact of the outflow coefficients, previously chosen in the range $0.1 < \rho_i < 0.7$. If instead all ρ_i were close to unity, this would mean that agents do not interact, but spend most of their money directly out of their own sector, like in the household case. In this case the sector response is simply a weighted mean of individual responses. On the other hand, if all

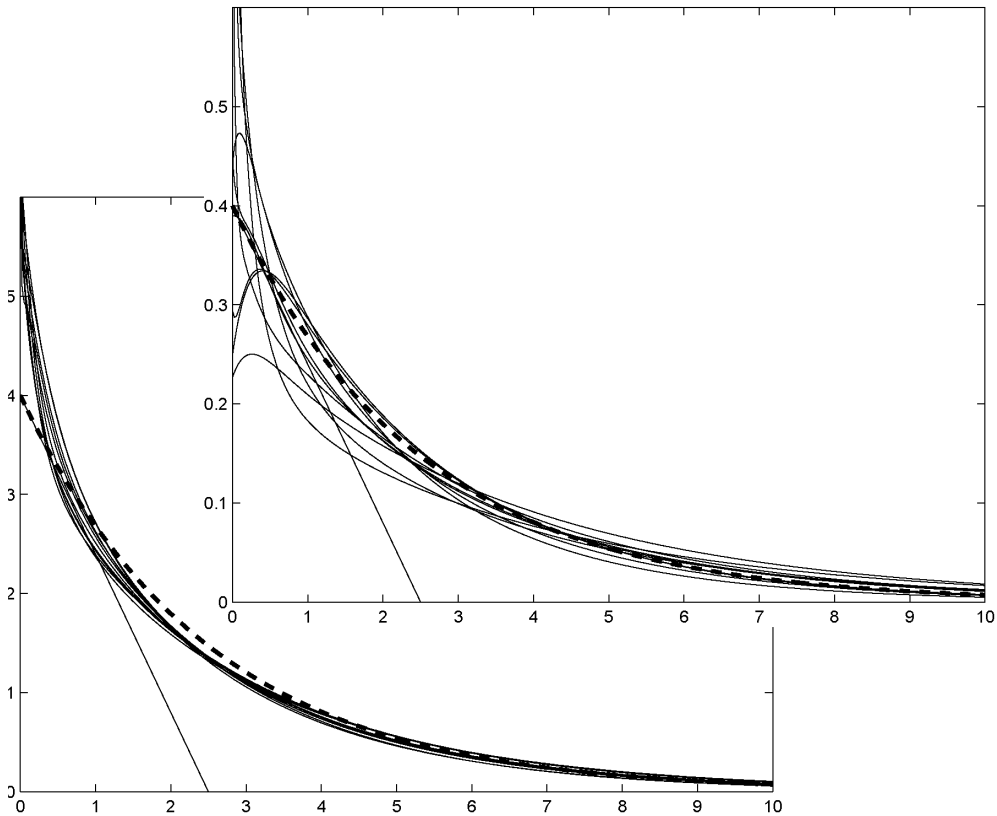


Figure 2.40: Aggregate responses for a 150-agent system (LL), and a 10-agent system (UR)

ρ_i were close to zero, this means that a unit of money in an average sense has to pass by many agents before it is spent out of the sector. We have simulated such a case, with $0.05 < \rho_i < 0.15$, i.e. a mean of 0.1. Figure 2.41 shows ten Monte Carlo responses for a 150-agent system and the predicted first-order response, which now is $h(t) = 0.1e^{-0.1t}$. In this case the time lag is 10 for the aggregated system, in accordance with (2.75). Compared to figure 2.40, lower left window, we see that

- the responses are closer to the proposed first-order approximation.
- the initial spikes, which are due to zeroes in transfer functions for individual agents in the network, still leave a mark on the aggregate

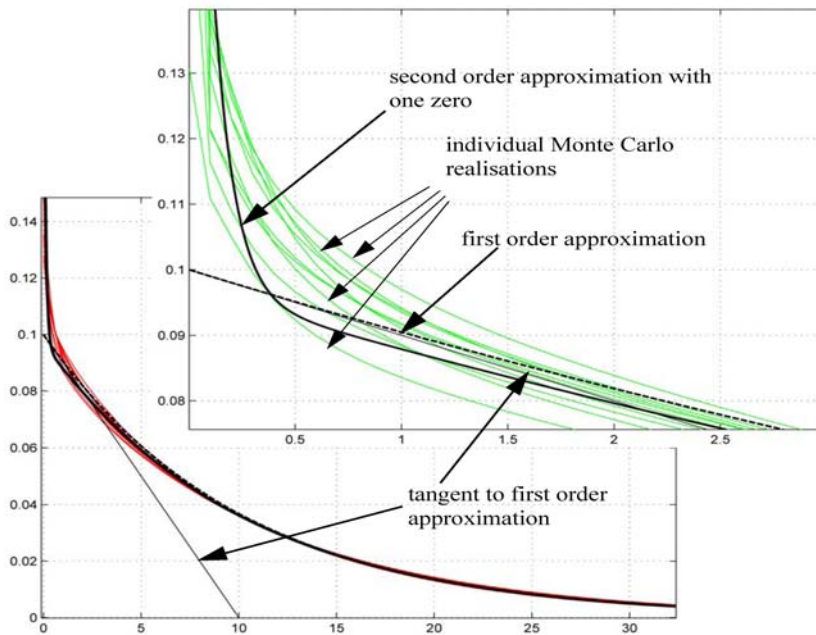


Figure 2.41: Ten aggregate responses for a 150-agent system with low ρ_i 's

response in the form of a corresponding initial peak, but a smaller and narrower one.

We will now try to explain this by exploiting an intermediate result from the proof given for Theorem 2.1. We again make the simplifying assumption that all agent transfer functions and all outflow coefficients ρ_i are identical. But now the agent transfer functions are identical, of the type (2.87). We divide by α_1 in the numerator and denominator, and get

$$h(s) = \frac{1 + T_z s}{1 + a_1 s + a_2 s^2} = \frac{n(s)}{d(s)} \quad (2.95)$$

The intermediate result (2.78) may be generalised to

$$h_a = \rho \frac{n}{d} + \frac{n}{d} (1 - \rho) h_a \quad (2.96)$$

(here dependence on s is omitted). Solving for h_a and using (2.95), we may write

$$h_a(s) = \frac{1 + \rho T_z \left(\frac{s}{\rho}\right)}{1 + (a_1 - T_z [1 - \rho]) \left(\frac{s}{\rho}\right) + \rho a_2 \left(\frac{s}{\rho}\right)^2} \quad (2.97)$$

We compare this to the first order approximation, which, following (2.75) and (2.17), is

$$h_a(s) = \frac{1}{1 + (a_1 - T_z) \left(\frac{s}{\rho} \right)} \quad (2.98)$$

Eq. (2.97) confirms that the zero will make itself felt²⁶ also for the aggregate system, as already observed through the Monte Carlo runs shown in figure 2.41. But (2.97) also tells us that its relative influence on system dynamics is less when ρ is small, which is supported by a comparison of figures 2.40 and 2.41. Furthermore, when ρ is reduced, the influence on system dynamics of the second-order term in the denominator in (2.97) decreases in relation to the first-order term. This also supports the observation that the system response is closer to that of the first-order approximation (2.98) when ρ is small.

2.9.2 Networks of units with irrational transfer functions

We will do an additional exercise to test the generality of the decaying exponential as a reasonable approximation to the response of an arbitrary network of linear units. We revert to the assumption in section 2.8 that all networks units are identical, so the simulation will be deterministic, not Monte Carlo. But now the dynamics of each unit are characterised by a rectangular impulse response, as indicated in figure 2.42. For comparison is shown the impulse response of the transfer function $1/(1 + \tau s)$ which we employed in connection with the derivation of the aggregation theorem. Both responses are PUA and have the same time lag.

Will an infinite network of such identical rectangular response units exhibit the same approximate exponential dynamics that we have seen till now? This cannot be solved algebraically, since these units (and therefore also the network consisting of them) cannot be described by differential equations (even if the system is still linear). But we can do this numerically, by employing the structure from the alternative proof of the aggregation theorem, see figure 2.35. The two blocks are now of the rectangular response type. Simulation is done in Simulink, where the block diagram looks like in figure 2.43. The two identical rectangular response units are indicated with

²⁶A response with a sharp initial spike in addition to a correspondingly smaller (so that the total is still PUA) exponential tail indicates that the transfer function $h(s)$ is close to not being strictly proper, i.e. the numerator having the same order as the denominator. This is a system where part of the input is (nearly) directly transmitted to the output. An economic unit with such a response spends a large part of received money immediately on receipt.

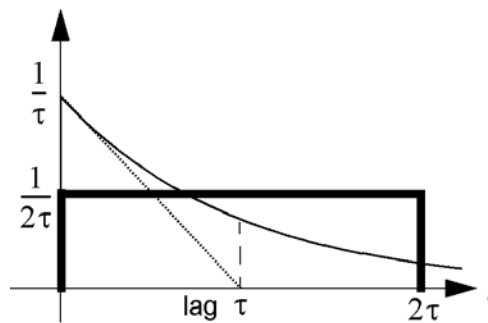


Figure 2.42: Rectangular PUA impulse response, compared to stable exponential

dashed outlines. The entity “tdel” is the length of the pulse, so that $\text{tdel} = 2\tau$. In the lower part of the diagram is a representation of the corresponding aggregate first order time lag. Its response is simulated and logged for comparison. Simulation results are shown in figure 2.44 for three values of the outflow coefficient: $\rho = 0.5, 0.15, 0.05$. We again observe the decaying exponential shape, except for an initial transient phase. The explanation for the strongly non-exponential graph for $\rho = 0.5$ is that when “leakage” out of the sector (network) from each unit is strong (ρ closer to unity), the non-exponential character of each unit’s dynamics will be reflected also in the network’s dynamics. A small ρ gives a much more exponential-like response. This is also what we observed in subsection 2.9 where the network consisted of widely different second-order rational transfer function units.

Note the time scales for the three responses. For a sector consisting of these “far-from-exponential” units we still have that the lag of the network’s impulse response is τ/ρ . This holds exactly when all units are identical. It also holds exactly in the more general case where units are arbitrarily and structurally different from each other, as long as they are PUA and have the same time lag. This was demonstrated in connection with the derivation of the series and parallel rules for calculating network time lags – equations (2.20) and (2.28), respectively.

2.9.3 Conclusion

The first order time lag approximation for a sector is supported by the above experiments. It works best for sectors with strong interaction/detour effect (for instance firms, as opposed to households).

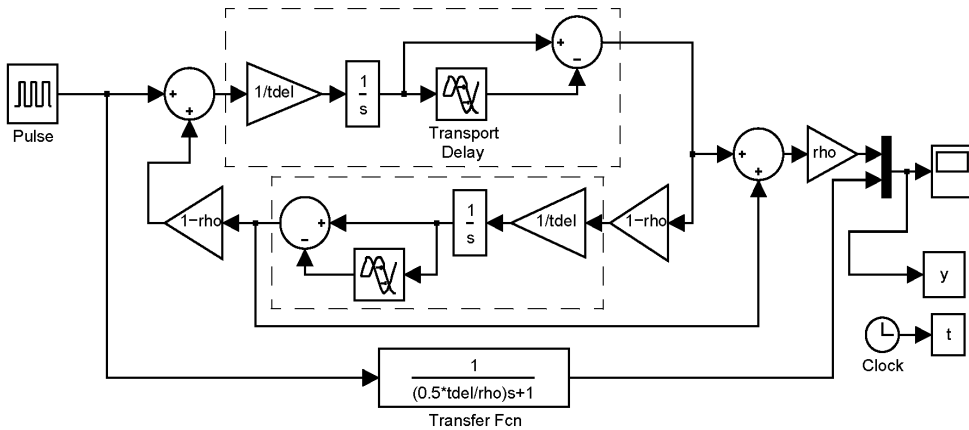


Figure 2.43: Simulink diagram with two rectangular impulse response blocks

Since individual agents' time lags and outflow coefficients are constant in each simulation, behavioural assumptions have been quite restrictive. But the model can easily incorporate time-varying parameters, so that for each agent, any of its coefficients may change at any time. The time lag of the sector will change accordingly. But decaying exponential dynamics will remain, and the sector time lag will be inversely proportional to (a somehow defined) mean ρ .

Parameters may also be dependent upon system states. In this case the model will become non-linear. We will see examples of this in chapter 6.

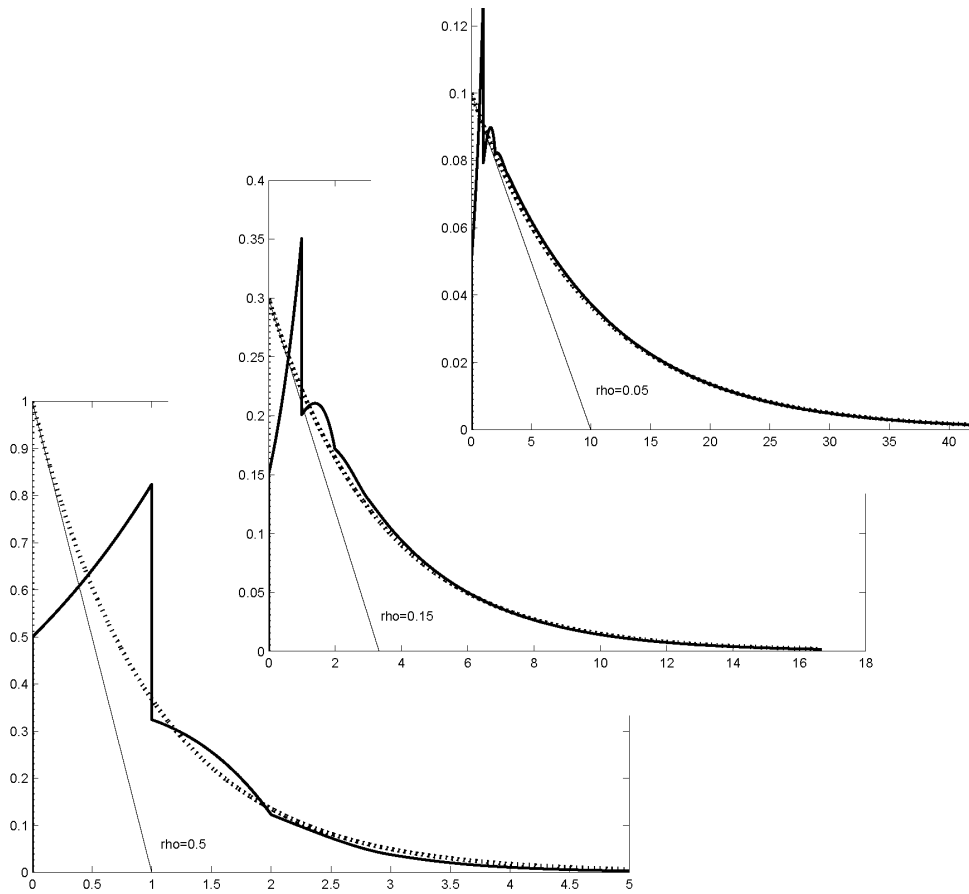


Figure 2.44: Three impulse responses for the system in figure 2.43, $\rho = 0.5, 0.15, 0.05$

Chapter 3

Discrete versus continuous time

We will in this chapter more comprehensively develop the argument that continuous-time modeling is superior to using discrete time. The analysis uses an example from Godley and Lavoie's book (Godley and Lavoie; 2007) that is in time-discrete format, and develops a continuous-time equivalent for the same model. This thesis is completely supportive of their stock-flow modeling paradigm. But the point of this chapter is to argue that continuous-time representation is superior for working with such models.

3.1 A time-discrete model

Figure 3.1 is a facsimile from table 3.4 in (Godley and Lavoie; 2007). It is the first and simplest of a series of discrete-time stock-flow macro models presented in their book.

We will use the symbols from figure 3.1, and the same values for parameters and variables, so that we may check whether we get the same time path as indicated in the columns 1, 2, 3, ∞ in the figure. As usual \$ is the generic monetary unit. "tu" is a generic time unit (but could be a quarter or a year). We have

G = government spending = constant = flow = 20 [\$/tu]. from period 2 onwards.

Y = gross income [\$/tu]

T = tax flow [\$/tu]

Period	1	2	3	∞
G	0	20	20	20
$Y = G + C$	0	38.5	47.9	100
$T = \theta \cdot Y$	0	7.7	9.6	20
$YD = Y - T$	0	30.8	38.3	80
$C = \alpha_1 \cdot YD + \alpha_2 \cdot H_{-1}$	0	18.5	27.9	80
$\Delta H_S = G - T$	0	12.3	10.4	0
$\Delta H_h = YD - C$	0	12.3	10.4	0
$H = \Delta H + H_{-1}$	0	12.3	22.7	80

Figure 3.1: A time-discrete G & L macro stock-flow model

YD = disposable income [$\$/tu$]

θ = taxation rate = 0.2 []

C = consumption flow [$\$/tu$]

α_1 and α_2 = coefficients, giving consumption out of disposable income and out of wealth (money stock) respectively; $\alpha_1 = 0.6$ [] and $\alpha_2 = 0.4$ []

H = current wealth (= money stock) [$\$$]. In this simple model, all money is held by households. H_{-1} is the value from the previous period, discrete time.

$\Delta H = H - H_{-1} = \Delta H_h = \Delta H_S$ in figure 3.1

All variables are zero before a government spending flow G is switched on at the start of period 2 and held constant during this period. This results in a gradual build-up of economic circulation and money stock H . As indicated in figure 3.1, variables converge towards an equilibrium after several periods.

We will discuss this model and its dynamics. First, we portray the model in the book as a discrete-time block diagram, figure 3.2.

The block containing z^{-1} is a backwards discrete time step operator, signifying the connection between H and H_{-1} . Ignore the dotted rectangle for the time being. The textbook way to approach the model is to partition the analysis into short-run "multiplier" dynamics, and (sometimes) long-run

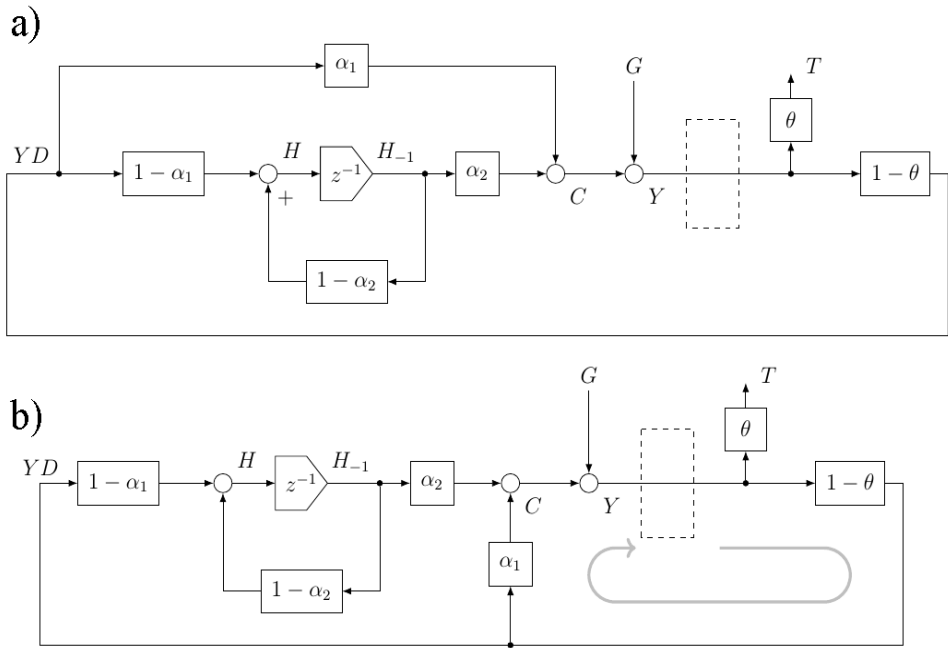


Figure 3.2: Godley & Lavoie model, diagrams are equivalent

dynamics. We will argue that this distinction is an unnecessary and artificial consequence of working in discrete time.

Consider the short-run situation (one period). We observe, especially from figure 3.2 b), that there is an infinitely fast loop (containing α_1 and indicated by the shaded closed arrow) in the system: Government spending G gives an immediate demand YD which immediately increases consumption C which, added to government spending G , immediately increases YD and so on. The mathematically correct way to approach this is to eliminate the instantaneous feedback loop by solving the set of equations:

$$YD = Y(1 - \theta) \tag{3.1a}$$

$$Y = G + C \tag{3.1b}$$

$$C = \alpha_1 YD + \alpha_2 H_{-1} \tag{3.1c}$$

Substituting (3.1c) in (3.1b), and inserting the result of this in (3.1a), and then solving for YD , we get

$$YD = \frac{1 - \theta}{1 - (1 - \theta)\alpha_1} G + \frac{(1 - \theta)\alpha_2}{1 - (1 - \theta)\alpha_1} H_{-1} \tag{3.2}$$

YD may alternatively be found numerically by iteration, but this is just a solution technique for a set of static algebraic equations and has nothing to do with dynamics (and iteration was not needed in this case since a closed algebraic result was easily found). But G & L choose to consider this to be a process that unfolds via a "multiplier" *within* the period, on p.69:

The initial \$20 government injection thus has ripple effects throughout the economy. The government injection has a multiple effect on income. This is the well-known Keynesian multiplier process, to be found in all elementary macroeconomics textbooks. Because perfect foresight has been assumed, households must know precisely how much will be produced and how much income the initial injection of government expenditures is able to generate. They must also know with certainty the various parameters of the overall economy (the average tax rate and the average propensity to consume out of disposable income) and they must know the initial injection. Here, all these multiplier effects are assumed to take place within the single period. Starting with no economic activity at all in period 1, the government expenditures taking place at the beginning of period 2 along with the standard multiplier process will bring about the numbers given at the end of period 2, as shown in Table 3.4. (*– in the book, facsimile in figure 3.1 above.*)

The assumption that agents think forward and proceed in lockstep over some agreed upon period, is not credible. And as long as all variables are assumed constant within a period, there are no dynamics there, only the solution to the equation set (3.1).

In addition to being a bad model, it also unnecessarily complicates the issue. If one assumes that there are dynamics unfolding within a "period", then that period is simply chosen too large. It should then be made so short that there is no need to assume movement within it. So criticism of this should not only, as Godley and Lavoie do on pp. 69–71, be that short-term equilibrium is not enough because one has to check out the long term. A more correct and stronger criticism should be that *there simply is no "short-term equilibrium"*; the concept is meaningless because it is just an artifice of using discrete time.

We will now portray the model given by (3.2) in a block diagram, see figure 3.3. The instantaneous feedback loop is now eliminated, but at a price: the explanatory power of the diagram (figure 3.2) in showing connections and causalities between key variables, is lost.

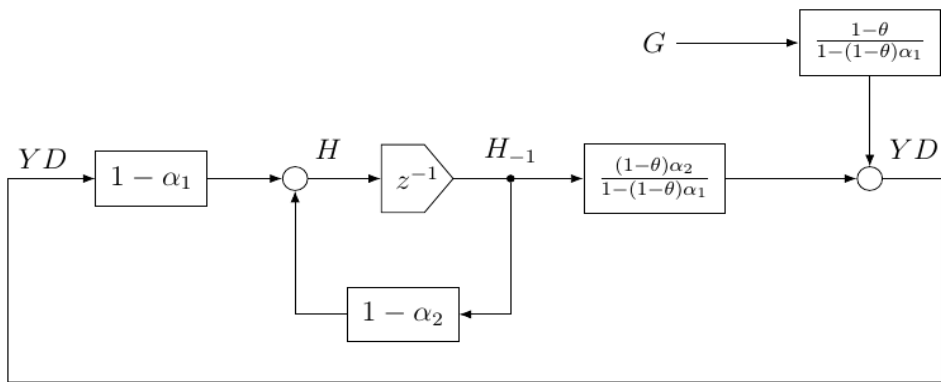


Figure 3.3: Block diagram with instantaneous loop resolved

Before proceeding with a time-continuous model of the same system, we will find the difference equation governing the dynamics of this first order discrete model. It is the equation for update of the sole system state, money stock H . From two equations in figure 3.1 we have $H = H + G - T$. We also have $T = \theta Y = YD \theta / (1 - \theta)$. Using (3.2) for YD and solving for H , we get

$$H = \frac{(1 - \alpha_1)(1 - \theta)}{1 - (1 - \theta)\alpha_1} G + \frac{1 - (1 - \theta)\alpha_1 - \theta\alpha_2}{1 - (1 - \theta)\alpha_1} H_{-1} \quad (3.3)$$

which is also given on p. 87 in the book.

3.1.1 A time-continuous and better model

We will now present a model in continuous time that closely emulates the behaviour of the original model, but which in many ways is superior (as will be demonstrated). We define continuous time to run in the same units (called "tu") as used in the discrete model. This simplifies comparisons, since all flows and parameters will then have the same numerical values (while H – since it is a stock and not a flow – will always be numerically the same regardless of time units employed). The time-continuous model is given in figure 3.4. This model is faithful to the original structure of the discrete G & L model. It embodies the same detailed information as in figure 3.2. We will see that it gives (nearly) the same transient behaviour and tends to precisely the same equilibrium values.

The model avoids infinite-speed money circulation (which does not occur in the real world) by introducing a very small time lag εT in the upper

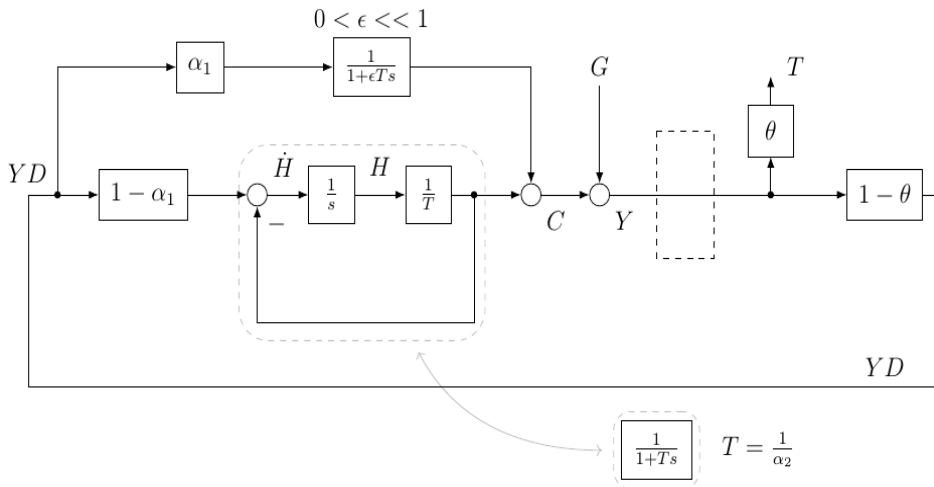


Figure 3.4: Continuous-time Godley-Lavoie model

branch. And in continuous-time simulation, the software allows us to choose – or it automatically chooses – an even smaller numerical simulation time step $\delta \ll \epsilon T$. This time step can be changed without impacting any parameter or variable values in the model.

Another advantage is that we can tweak the time constants T and ϵT without this changing anything else in the system's equations. This is in contrast to a discrete model, where the time step is always dimensionless "unity" (and often too large in relation to possible fast modes¹ of the system). Also, any change in the system's time lag(s) or the discrete time step require changes in most of the numerical values of the discrete model's parameters, and its input value(s).

The model in figure 3.4 is second-order, so in that sense one pays a price compared to the simpler first-order discrete model. But with today's simulation software this is no problem. The dots in figure 3.5 are the responses of the G&L model (as given in figure 3.1). Comparing these to the corresponding continuous responses (of the model in figure 3.4), we observe a

¹In dynamic state space terminology, a *mode* is a component of system states with dynamics related to a corresponding eigenvalue of the system. The simulation time step should as a general rule be chosen smaller than the inverse of the largest absolute value among the eigenvalues, corresponding to the fastest mode. In discrete models, one doesn't have this freedom of choice, but is stuck with a simulation step size identical to the discrete time interval chosen.

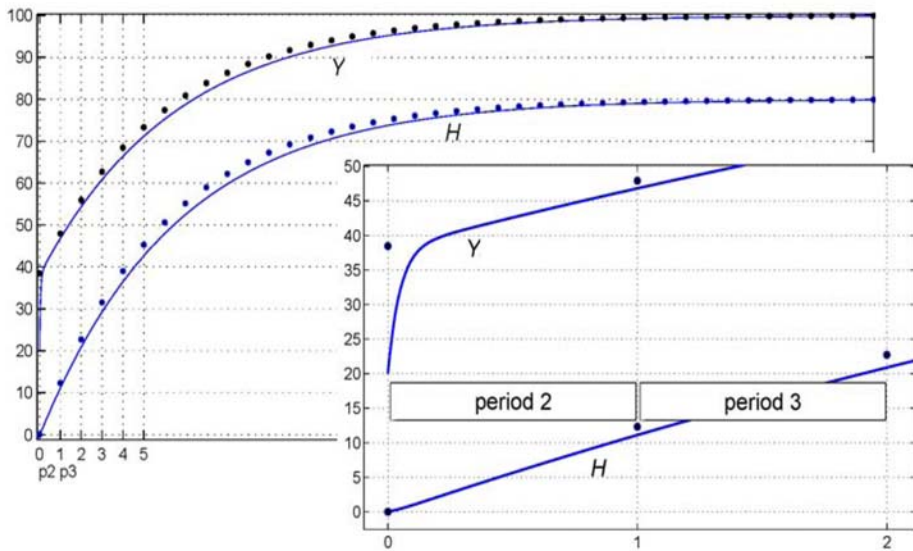


Figure 3.5: Discrete-time and continuous-time responses

close resemblance. The lower right graphs are magnifications of the first two time periods.

Note that the time periods in figure 3.1 are shifted compared to the continuous time axis: The discrete G input is switched on at the start of "period 2", while the corresponding continuous G is switched on at time zero. Another possible source of confusion is that in the table in figure 3.1, the columns for periods 2 and 3 have the system state H positioned (seemingly) synchronously with the G input. But the discrete dynamics are such that H changes *after* the G step has made itself felt. So column 2 in figure 3.1 must be understood such that the H value of 12.3 is associated with *the end* of period 2 while the G step of 20 in the same column occurred at *the start* of period 2. Only with this interpretation may both be associated with period 2 as done in figure 3.1. This is also emphasised in the book, p. 69.

We note that the continuous-time responses converge towards the same equilibrium values as in the discrete case. The initial jump in YD is also close to the corresponding discrete jump, see the magnified lower right graph in figure 3.5. The slightly higher time lag for the continuous version is mainly a consequence of the upper branch for household spending not being fully instantaneous. The household transfer function from YD to C is of the type indicated in figure 2.41 (with $\rho = 1$), and by eq. (2.97) – a transfer function

consequences. For every extra lag introduced (and each lag implies a money buffer in the same location/sector), the system's order increases by one, but as already stated this is no practical obstacle with modern simulation software. Figure 3.7 shows a possible version. We have now introduced a

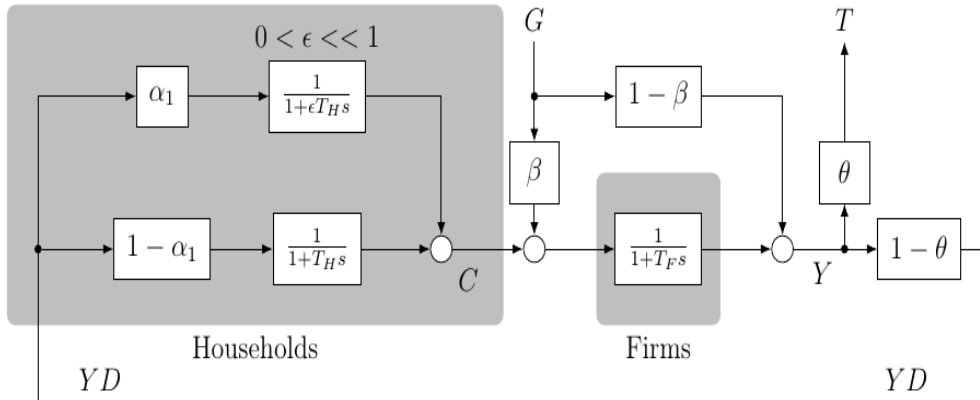


Figure 3.7: Continuous model with a firm sector and "split" government spending

sector of firms, corresponding to the dotted square suggested in the earlier block diagrams in this chapter. This implies more realism in the model: households and the government buys from a firm sector, and households are paid income by the government and firms. We have split the government spending flow into a wage flow to pensioners and households employed in the public sector, and a flow directly to the firm sector, accounting for government purchases. With two sectors, households and firms, the model resembles the one introduced in section 2.7.

Compared to the adjustment speed of the model in figure 3.4, observed in figure 3.5, dynamics will be much slower. The circular money flow is additionally lagged when passing through the firm sector, and more so there than in the household sector, as already argued in section 2.8. Again we wish to calculate the time lag from G to Y . We build upon figure 3.6, and extend it with the firm block connected in series and also in the feedback loop, see figure 3.8. This block diagram is the same as figure 3.7, slightly manipulated, and with $T = T_H = 1/\alpha_2$. We use the same procedure as above to find the time lag from G to Y . The lag of the lower feedback

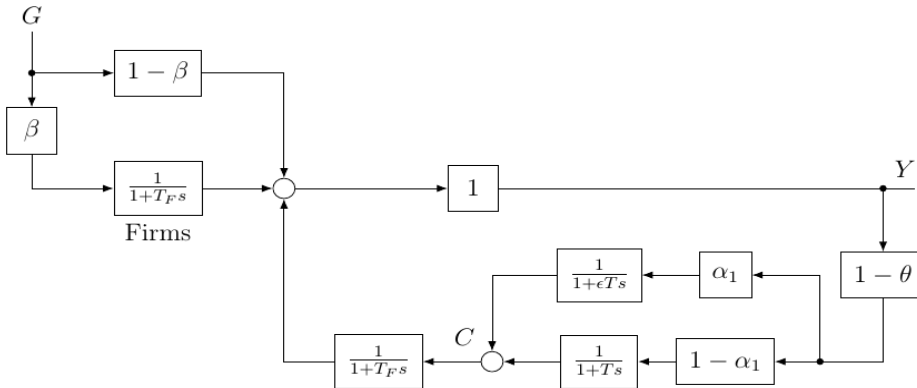


Figure 3.8: Slower dynamics when adding a firm sector to the model.

branch is, using (3.4) and the series rule:

$$\bar{T}_{fb} = \bar{T}_H + T_F = \frac{(1 - \alpha_1)}{\alpha_2} + T_F, \quad (3.6)$$

again simplifying by exploiting $\varepsilon \approx 0$. Using the parallel rule on the two branches containing β , then the series rule, then using \bar{T}_{fb} for T_2 and θ instead of ρ in (2.30), we get

$$\bar{T}_{YG,2} = (0 + \beta T_F) + \frac{0 + (1 - \theta)\bar{T}_{fb}}{\theta} \quad (3.7)$$

Choosing numerical values $\beta = 0.2$ and $T_F = 3T_H = 3/\alpha_2$ gives $\bar{T}_{YG,2} = 35.5$, a much slower adjustment process than the earlier process with lag $\bar{T}_{YG} = 4$, as expected. Responses to the step function $G = 20$ are shown in figure 3.9. The lag of $Y(t)$ is indicated by a line in the figure. The system is fourth order, so the line is not the precise tangent to the graph at the origin. But it gives a very good indication of adjustment speed. It does not intersect with the horizontal asymptote at $\bar{T}_{YG,2} = 35.5$, but at a somewhat higher value. We will find this value – doing an additional exercise in model reduction and in applying the time lag rules developed earlier. We seek a good first order approximation to the transfer function from G to Y . It must be of the form

$$h_{YG}(s) = \frac{1 + T_1 s}{\theta(1 + T_2 s)}, \quad (3.8)$$

which gives $Y(\infty) = 100$ for $G = 20$ as required. $T_2 > \bar{T}_{YG,2}$ is the time of "tangent" intersection with the $Y(\infty)$ asymptote that we are searching.

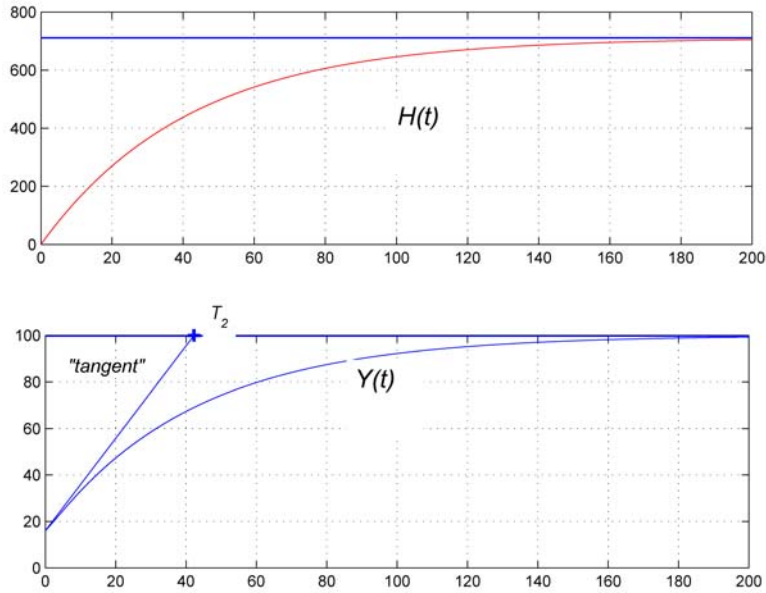


Figure 3.9: Output Y and money stock H for system with firm sector.

And with a $T_1 > 0$ in h_{YG} , the step response starts above zero, at $\frac{GT_1}{\theta T_2}$ (this follows from using the Initial Value Theorem² for Laplace transforms). From figure 3.8 we observe that there is a direct contribution to Y at $t = 0$ of value $G(1 - \beta)$, giving the size of the initial jump. We then have

$$(1 - \beta) = \frac{T_1}{\theta T_2} \quad (3.9)$$

We also have, using (2.17),

$$\bar{T}_{YG,2} = T_2 - T_1 \quad (3.10)$$

These two equations may be solved for T_1 and T_2 , giving

$$T_2 = \frac{\bar{T}_{YG,2}}{1 - (1 - \beta)\theta} = \frac{35.5}{0.84} = 42.26 \quad (3.11)$$

Concerning the long-term equilibrium, we get $Y(\infty) = 100$ as before. This is to be expected since the system still satisfies $Y(\infty) = G/\theta$. There

²The initial value theorem is $\lim_{t \rightarrow 0} f(t) = \lim_{s \rightarrow \infty} s f(s)$

are now, however, two equilibrium states (buffers) in the system, in the household and the firm sector. Their sum $H(t)$ is shown in the upper part of the figure. Corresponding to the much larger time lag, the amount of money in circulation when equilibrium has been reached, is also much larger, cf. (2.37).

Generally, splitting up flow paths and/or insertion of additional time lags anywhere in the network, as long as the associated transfer functions are, or remain, PUA, will not change $Y(\infty)$. One could, for example, detail the system further by partitioning the firm sector into an investment/intermediate goods sector and a consumption goods sector, like in subsection 2.8.3, without impacting $Y(\infty)$.

By the above, the superiority of working in continuous time should hopefully have been demonstrated.

Chapter 4

Lending and financial accumulation

Conventional macroeconomic modeling has to a large degree abstained from incorporating money and debt. Here is a "mea culpa" about ignoring debt – inspired by the global financial crisis from approximately 2008 onwards:

If there is a single word that appears most frequently in discussions of the economic problems now afflicting both the United States and Europe, that word is surely debt.

...

Given the prominence of debt in popular discussion of our current economic difficulties and the long tradition of invoking debt as a key factor in major economic contractions, one might have expected debt to be at the heart of most mainstream macroeconomic models, especially the analysis of monetary and fiscal policy. Perhaps somewhat surprisingly, however, it is quite common to abstract altogether from this feature of the economy. Even economists trying to analyze the problems of monetary and fiscal policy at the zero lower bound – and yes, that includes the present authors [references not listed in this thesis]¹ – have often adopted representative agent models in which everyone is alike and the shock that pushes the economy into a situation in which even a zero interest rate is not low enough takes the form of a shift in everyone's preferences. Now, this assumed preference

¹We use [brackets] to indicate when quotations contain references which are not listed in the bibliography of this thesis.

shift can be viewed as a proxy for a more realistic but harder to model shock involving debt and forced deleveraging. (Eggertsson and Krugman; 2012)

The next three chapters (4, 5, 6) of this thesis will extensively try to incorporate debt, its dynamics and crisis mechanisms based on the already developed stock-flow toolbox. We will add a financial sector in our stock-flow structure in this chapter.

Outflows from units discussed until now have been such that they are solely controlled by the unit or sector of units, with their time lag T as the crucial parameter. But as earlier mentioned, there are outflows that are not under the control of the agent/sector in question, but imposed on them from other parts of the economy – *non-discretionary* flows. A crucial assumption in later chapters is that non-discretionary flows in the form of debt service may become too large in relation to discretionary flows in the macroeconomic monetary network, and this will lead to crisis.

4.1 The basic moneylending structure

In this section we will start by examining lending and debt dynamics based on a simple "classic moneylender" model: The lender lends, and based on this, receives a stream of money (think: physical cash, "gold coins") which is the sum of interest and repayment (= debt service). The lender can then choose to recycle a share of this flow as new loans, and spend the rest. The coefficient that decides the recycling share is called σ , with $\sigma < 1$. Obviously the size of σ is important for whether the aggregate of loans will grow or not. This model is too simple for a reasonable description of how banks behave in a modern framework. We will discuss that in the next chapters, but will observe that the basic compound interest mechanism is at work there, too. Furthermore, the moneylender model to be presented here correctly describes non-bank lending, which – as we will see later – constitutes the most serious factor for development of debt-related crisis.

Debt service, like taxes, is a non-discretionary flow. Taxes are dynamically trivial since a taxation event (in continuous time expressed as an impulse function) imposed at some moment, implies no future related flows. Debt service flows however, has interesting dynamics that unfold over time: the initial one-shot input (received loan) leads to a stream of future events (debt service). A loan may in our terms be considered an impulse input to a unit, and then the debt service flow becomes the impulse reponse of a *debt service subsystem* (from now on abbreviated "DSS"). Seen from the

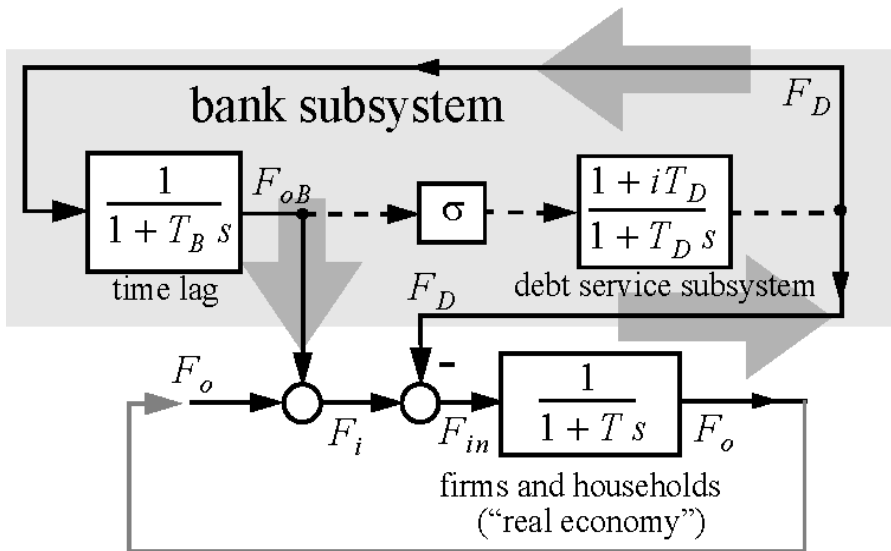


Figure 4.1: A bank subsystem with recycling of loans

debtor's position this is a non-discretionary flow. The model may be explained via the block diagram in figure 4.1. In the lower part of the figure is a debtor unit, which may be a single agent or a sector as discussed earlier. If the unit is the entire aggregate of firms and households, F_o is recycled to the input, indicated by the shaded arrow lowermost in the figure (then we will in between use the term "the real economy" – "RE" – for this unit in the following). The flow F_{in} is what is left over for wages and the purchase of goods and services after debt has been serviced and new loans have been received.

A new phenomenon is introduced in this block diagram: F_D is not the outflow from an agent or sector, but the result of a *rule* (the loan contract). This rule decides a (non-discretionary) flow F_D , which is subtracted at the real economy input and inserted as an inflow to the bank unit (which is the aggregate of all banks if the debtor unit is an aggregate). By this the accounting remains correct: money removed from one flow is input somewhere else. Even if F_D is paid out by the unit, it is subtracted at the *input* not the output, since F_D is non-discretionary. The output consists only of such flows that the unit can control itself, i.e. discretionary flows.

To indicate the presence of rule-based interaction as opposed to interaction flows due to "vessel dynamics" only, the corresponding lines are dotted

in the figure. The use of arrows to indicate "influence" in a block diagram represents a break with the additional convenient symbolism of arrows in these diagrams until now, where they have been synonymous with money flows. The actual flow direction of money in the figure is indicated with thick shadowed arrows.

We have here assumed a scenario where the flow of new loans is a strict feedback from what banks receive in debt service on current loans; a share σ is re-lent. This is pure lender-controlled financial accumulation. Later we will discuss cases where the flow of new loans is not directly a result of what inflows banks receive. We will use the term "bank" here in a generic sense: any type of unit that has a financial claim (here called a "loan") on another unit as long as the claim implies a contracted future stream of returns. We will call σ the *financial reinvestment coefficient*, abbreviated *FRC*².

The interest rate is i and duration of loans is T_D . Debt service is modeled here as a continuous flow, while in the real world debt is serviced in time discrete packets. In our continuous-time setting this could have been precisely accounted for by a train of impulse functions, but this is not necessary, following the earlier given argument in subsection 2.3, in connection with figure 2.4.

The model in figure 4.1 has a great advantage: It allows for calculating the dynamics of an aggregate economy where current debt service is used continuously to extend new loans, and where both the effects of interest rate and loan duration is accounted for. This is in contrast to much of the Post Keynesian and Circuitist literature, where one often – because of the tools used – has to abstract from interest and also assume that loan extension and repayment takes place in distinct and synchronised "rounds".

The (aggregate) "bank" in the figure is modeled as a first order linear system, assuming that the flow received by the bank is output again, dispersed with some lag T_B . This implies that the money held by the bank³ is $M_B = F_{oB}T_B$. The outflow F_{oB} consists of both the bank's paying for expenses, and its new loans flow which is a share σ of F_{oB} .

It now remains to explain the DSS in the figure. The transfer function

²It may be considered a "savings coefficient" for the bank. In this thesis we use the greek σ instead of the usual s for "savings coefficient". The symbol " s " is reserved for the Laplace transform free variable or the differentiation operator. See also appendix A.

³Some readers may object that the concept of banks "holding money" is meaningless, since banks may be considered to create money when lending, and destroy money when loans are repaid. This is the Post Keynesian position, which this author supports. We will treat this topic in chapter 5. But at this stage, it is for purposes of simplified presentation convenient to assume that the bank works like a non-bank financial institution (from now on, just "non-bank"), in the sense that it does not create or destroy money.

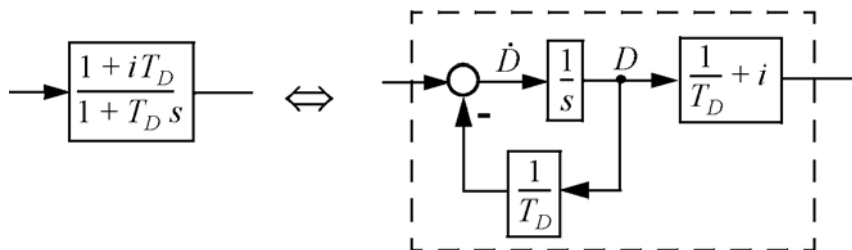


Figure 4.2: Equivalent debt service block

is

$$h_D(s) = \frac{1 + iT_D}{1 + T_D s} \quad (4.1)$$

which may be explained by introducing the equivalent "exploded" structure shown to the right in figure 4.2. Now debt D is visible in the subsystem. This DSS, with inflows and outflows as in figure 4.1, corresponds to the equations

$$\dot{D} = \sigma F_{oB} - \frac{1}{T_D} D \quad (4.2a)$$

$$F_D = \left(i + \frac{1}{T_D}\right) D \quad (4.2b)$$

This scheme (from now on called the "*exponential debt service scheme*") is unconventional, since both the principal and interest component is proportional to remaining debt. This differs from for instance an annuity scheme where the sum of principal and interest is constant, or a bond-type scheme where principal is only paid (in its entirety) when the loan matures. The advantage of the scheme (4.2) is that it allows for analysis using eigenvalues, and finding algebraic solutions – while annuity or bond-type dynamics involve time delays and are therefore algebraically not tractable. It will, however, be demonstrated that differences in total system behaviour are unimportant in regard to which scheme is assumed. Figure 4.3 shows the debt service flows for the exponential debt service scheme compared to the annuity-type scheme. If we consider a loan of 1 \$ extended at $t = 0$, these debt service flows are the impulse responses of the debt service subsystems. For approximate equivalence, we suggest that both types of DSS should have the same mean lag. This means that loan durations differ, with $T_{D2} = 2T_{D1}$ (this multiplicative factor will be somewhat adjusted below, in subsection

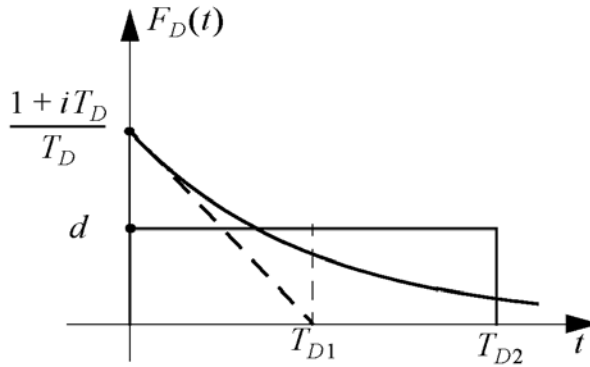


Figure 4.3: Debt service for annuity and exponential schemes

4.3.2). Mathematically, the duration of the exponential debt service scheme is infinite, but we define it to be T_{D1} , since this is the mean lag of the graph. The areas under the graphs correspond to the accumulated debt service sums. They are > 1 , so the DSS is not PUA (it would have been if i was 0)⁴. The value of the constant parameter d in the figure, which gives the annuity debt service flow, is derived below.

4.2 The annuity-type debt service subsystem

We assume that a loan of 1 \$ is extended at $t = 0$, and demand that the discounted value of a received constant flow d between 0 and T_D shall be equal to 1:

$$d \int_0^{T_D} e^{-it} dt = 1, \text{ which gives } d = \frac{i}{1 - e^{-iT_D}} \quad (4.3)$$

If the loan is a perpetuity i.e. $T_D = \infty$, (4.3) gives $d = i$ as expected. For the special case $i = 0$, L'Hopital's rule, or the integral in (4.3), gives $d = 1/T_D$, also as expected.

We may now construct a subsystem for this annuity type DSS, that has a rectangular impulse response with amplitude d . It is shown in figure 4.4. The subsystem works like this: A new loan (an impulse) is received, and the

⁴Note at this stage that the rules for model reduction to a first order model derived in section 2.6 do not allow debt service subsystems in the network to be included. A DSS is not PUA, and the rule-based outflow implies a type of connection with the rest of the network that does not allow such simplification.

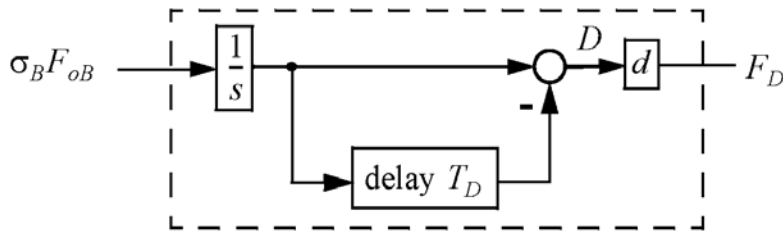


Figure 4.4: Debt service subsystem with annuity scheme

integrator makes the value in the upper branch jump to the size of the loan and stay there. After a duration of T_D , the lower branch jumps to the same level and is subtracted from the upper branch value. This ensures that debt service for that particular loan stops when the loan terminates. We have a rectangular response with amplitude equal to the size of the loan, multiplied with the factor d to give the correct debt service outflow from the DSS.

This DSS contains a time delay, and algebraic solution of systems containing time delays is generally not possible. But the system is still linear. Therefore a continuous inflow of new loans will, by convolution with the DSS impulse response, still give the precise debt service outflow. In other words: the effect of continuous recirculation of loans in a macroeconomic model may be correctly accounted for also in the annuity case. And we will see below that in this special case, stability may be verified algebraically in spite of eigenvalues not being available.

4.3 When may debt “explode”?

A widely covered topic in the literature and a persistent political-economic, moral and religious issue since ancient times is the mechanism of lenders accumulating financial claims on the rest of society by re-lending income from current loans. The problem of compound interest leading to exponential debt growth is recognised for instance in the Bible, where a “jubilee” is proscribed every 50th year to reset outstanding debt to zero (see appendix B). A contemporary historical treatment of debt and corresponding accumulation is (Hudson; 2000).

Obviously, a persistent re-lending of financial income *may* lead to financial debt/asset polarisation in a society. The structure in figure 4.1 allows us to check the conditions for this to occur with our model. Debt/asset polarisation corresponds to instability of this linear system. If we initially confine

ourselves to a system with an exponential debt service scheme, stability may be checked by considering system eigenvalues. By inspection of figure 4.1, we see that system dynamics are decided entirely by the shaded “bank” part of the structure. The dynamics of the lower “debtor” part do not feed back to the bank part and is therefore decided solely by what happens there. The characteristic equation for the bank part is

$$(1 + T_B\lambda)(1 + T_D\lambda) - \sigma(1 + iT_D) = a_2\lambda^2 + a_1\lambda + a_0 = 0 \quad (4.4)$$

A necessary (and for a second order system like this also sufficient) condition for the system’s two eigenvalues to be negative (i.e. stable system) is that all coefficients a_k in the characteristic polynomial have the same sign. a_1 and a_2 are always positive, while $a_0 = 1 - \sigma(1 + iT_D)$ may be < 0 for certain parameter values. Then one eigenvalue is in the right half plane. We have instability (= debt growth = financial accumulation). $a_0 < 0$ – which gives debt growth – corresponds to:

$$\sigma > \frac{1}{(1 + iT_D)}, \text{ or equivalently:} \quad (4.5a)$$

$$i\sigma > \frac{1}{T_D}(1 - \sigma), \text{ or} \quad (4.5b)$$

$$iT_D > \frac{(1 - \sigma)}{\sigma} \quad (4.5c)$$

We note that T_B is not part of the instability condition. If the condition (4.5) is fulfilled, debt growth is exponential (abstracting from an initial transient component due to the other, stable eigenvalue). Loan duration T_D may be in the order of – say – a decade. The bank time lag T_B should realistically be in the weeks/months range. So we may assume $T_B \ll T_D$. Following the argument in subsection 2.7.3, this means that the bank time lag subsystem in figure 4.1 may reasonably be substituted by unity. If we also ignore the debtor subsystem which has no impact on dynamic properties as already mentioned, the simplified remaining system needed to discuss debt build-up dynamics becomes as shown in the block diagram to the left in figure 4.5. To the right we have inserted the equivalent DSS from figure 4.2 so that the sole system state, D , is shown. This block diagram corresponds to the autonomous first order linear differential equation

$$\dot{D} = \left(-\frac{1}{T_D} + \sigma\left(\frac{1}{T_D} + i\right) \right) D = \lambda D \quad (4.6)$$

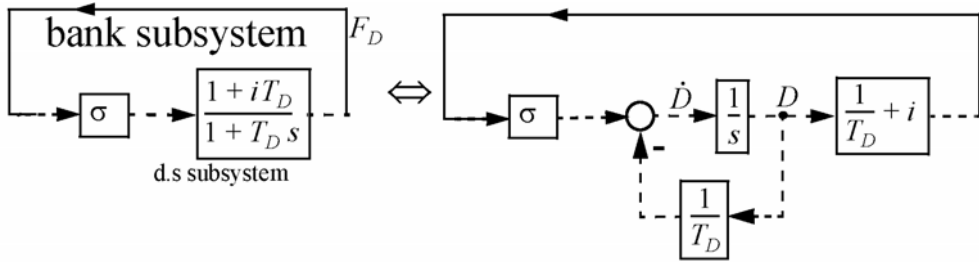


Figure 4.5: Simplified accumulation system

which has the solution $D = D_0 e^{\lambda t}$, where D_0 is initial debt. We have exponential growth for $\lambda > 0$, which corresponds to condition (4.5).

We will now discuss the roles of the three parameters T_D, i, σ : From (4.5c) we observe that a percentual increase in σ has a stronger effect towards accumulation than a similar increase in i . This may seem counter-intuitive, since the focus in this type of discourse is usually the impact of i .

For $\sigma = 1$, (4.6) becomes $\dot{D} = iD$, the “classic” differential equation for accumulation through compound interest, which will then take place for *any* $i > 0$. An expression of the fascination with – and alarm against – this phenomenon is the table in figure 4.6 which is a facsimile from (Kennedy; 1991). One pfennig (0.01 Deutsche Mark – this was written before the advent of the euro) deposited in year 0 at 5% interest is by 1990 worth 134 billion massive spheres of gold, each the size of the Earth. Admittedly, 5% is in real terms a fairly high interest rate, but the table still illustrates the dramatic dynamics of exponential (financial) growth⁵. In the following we will term such a runaway process *debt-asset polarisation*.

Another implication of (4.5) is that cet. par., a large T_D means steeper debt growth. If the loans are perpetuities ($T_D = \infty$), we have debt growth regardless of the size of σ and i , with $\dot{D} = i\sigma D$.

The same is the case if the lender relends all the repayment and spends only out of interest income (which will be a premise in chapter 5). Then we also get

$$\dot{D} = \left(-\frac{1}{T_D} + \left(\frac{1}{T_D} + \sigma i \right) \right) D = i\sigma D \tag{4.7}$$

⁵ Allegedly also commented by Albert Einstein: “the most powerful force in the universe is compound interest.” Ironically, this quotation is mostly used today not in the spirit of its critical originator: it is touted to market financial investment.

1 Danach ergeben sich vom Jahr 0 bis 1990 bei jährlichem Zinszuschlag folgende ausgewählte Kontostände:

Jahr	Rechnungseinheit DM	Rechnungseinheit kg Gold am 9. 1. 90 18 500,- DM/kg	Rechnungseinheit goldene Erdkugeln (5,973 E + 24 kg]
0	0,01		
95	1,03		
100	1,31		
142	10,20		
189	101,10		
236	1001,55		
296	18708,22	1	
438	19 094 706	1026	
1466	1,16 E+29	6,22 E+24	1
1749	1,148 E+35		1 Mill.
1890	1,116 E+38		1 Mrd.
1990	1,468 E+40	8,026 E+35	134 Mrd.

Heinrich Haußmann: *Der Josephspfennig*. Fürth 1990.

Figure 4.6: The dramatic dynamics of exponential growth

4.3.1 Pyramid game ("Ponzi") dynamics

The feedback model in figure 4.5 may be applied to a pyramid or Ponzi-type (Zuckoff; 2005) "investment" scheme⁶. The debt service flow plus Ponzi agent consumption must be matched by a flow of additional borrowing (from current and additional gullible investors). See figure 4.7, which – except for the added input of a required consumption flow – corresponds to figure 4.5 with $\sigma = 1$. Money flows are indicated with thick arrows. As pointed out above, $\sigma = 1$ always results in runaway exponential dynamics with growth rate i , regardless of values of T_D . An addition is that the Ponzi agent demands a consumption flow $C(t)$, which is an exogenous input in the model. The time path of $C(t)$ may be constant or increasing with time. As long as its increase rate is lower than that given by the endogenous dynamics due to the positive feedback from debt service to additional lending, it does

⁶ A very large such scheme was run by Bernard Madoff, New York. His promised return rate was $i = 15\%$. Madoff defrauded his clients of almost \$65 billion in the biggest Ponzi scheme in history. The operation was revealed to authorities in December 2008.

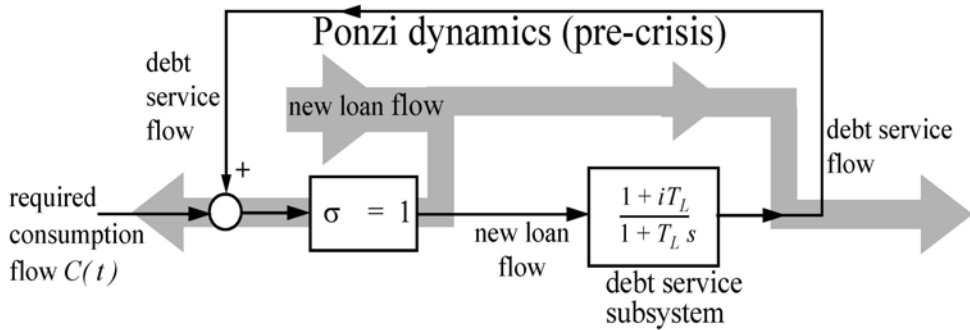


Figure 4.7: Ponzi (pyramid game) process

not impact much on the debt growth rate. Assume that the Ponzi agent demands a consumption flow

$$C(t) = e^{\beta t}, \text{ where } 0 \leq \beta < i \tag{4.8}$$

Without loss of generality (since this is a linear time invariant – "LTI" – system) we have chosen $C(0) = 1$ [\$/y], and we start the Ponzi process with no initial debt, $D(0) = D_0 = 0$ [\$]. Then the algebraic solution for the time path $D(t)$ is

$$D(t) = \frac{1 + iT_L}{T_L(i - \beta)} e^{it} (1 - e^{-(i-\beta)t}), \tag{4.9}$$

where the last exponential term $\rightarrow 0$ as $t \rightarrow \infty$.

Note that this simple linear model is only acceptable for the pre-crisis phase, where the Ponzi agent is able to borrow all that is needed to continue the process. A model that also accounts for the (sooner or later) collapse of the scheme could be developed by introducing nonlinearities and additional feedbacks in the system.

4.3.2 Accumulation with annuity-type debt service

We now want to check conditions for accumulation (instability) when the DSS is not of the exponential type as in figure 4.2, but of the annuity type, shown in figure 4.4. We also in this case choose to ignore the bank time lag subsystem, which is set to unity. The transfer function for the annuity DSS is

$$h(s) = \frac{d}{s} (1 - e^{-T_D s}) \tag{4.10}$$

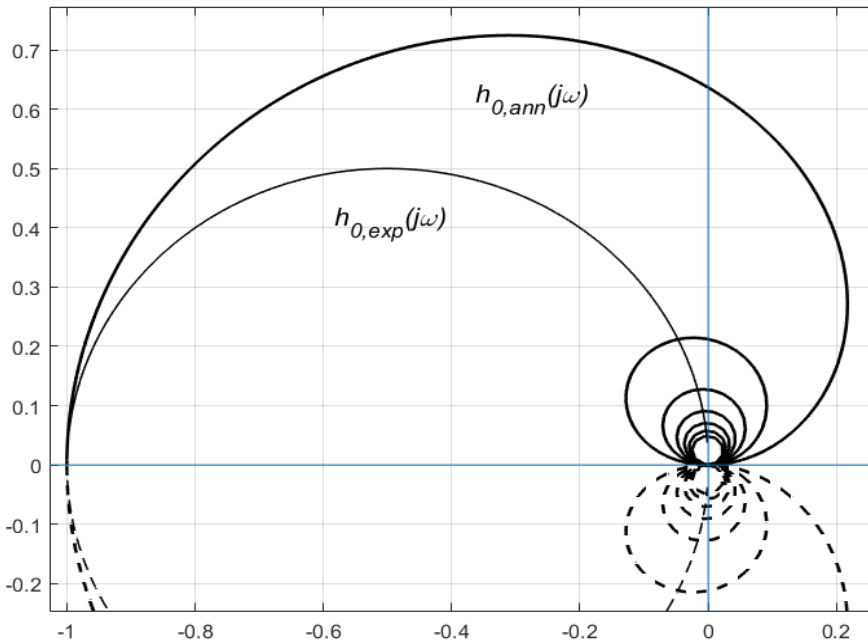


Figure 4.8: Polar diagrams of h_0 for stability check; annuity and exponential DSS

where d is given by (4.3), and T_D is the duration of the loan. When we close the loop, we don't obtain a characteristic polynomial but an irrational expression, due to the term $e^{-T_D s}$. Therefore we cannot check instability via eigenvalues. But since the system is still linear, we may use the Nyquist stability criterion (described in most control engineering textbooks). It is based on a frequency response polar plot of the loop transfer function $h_0(s)$, which in this case is

$$h_0(s) = -\sigma \frac{d}{s} (1 - e^{-T_D s}) \quad (4.11)$$

(a minus sign has to be used in h_0 because in Nyquist criterion the feedback must be negative, while the feedback is positive in our case.) The frequency response, given by setting $s = j\omega$ in $h_0(s)$, is displayed in a polar diagram⁷ in figure 4.8. When ω takes on values from $-\infty$ via 0 to ∞ , we get a corresponding closed graph for the frequency response $h_0(j\omega)$ as displayed

⁷In a polar diagram the length of the radius vector to a point on the graph of a complex-valued function $h_0(j\omega)$ is $|h_0(j\omega)|$, while the angle between the positive real axis and the radius vector is $\angle h_0(j\omega)$.

in the figure. The dotted half of the graph corresponds to $h_0(j\omega)$ for $\omega < 0$. When h_0 is open-loop stable, which it is in this case (since the impulse response goes to zero with increasing t), the Nyquist criterion simply says that the closed-loop system is stable when the leftmost part of the graph crosses the negative real axis to the right of the point -1 . The figure also shows the corresponding graph when the DSS is of the exponential type (where we used eigenvalues to check instability). The graph with this DSS is simply a circle, indicated with a thin line. In the figure, the choice of parameters σ, i, T_D is such that both graphs go precisely through -1 , which means that the two corresponding closed-loop systems are on the border of (in)stability. The chosen parameter values correspond to the two dots in figure 4.9 below.

While the Nyquist criterion as a general rule only gives a result based on a graph, in this special case we may employ it algebraically. If we consider (4.11) with $s = j\omega$, we see from angle and absolute value that the leftmost crossing of the negative real axis must take place for $\omega = 0$. We have

$$h_0(j0) = \lim_{\omega \rightarrow 0} \left(-\sigma \frac{d}{j\omega} (1 - e^{-j\omega T_D}) \right) = (\text{real}) = -\sigma T_D d \quad (4.12)$$

We substitute (4.3) for d . The Nyquist criterion, and (4.12) then gives the condition for financial accumulation:

$$\sigma > \frac{1 - e^{-iT_D}}{iT_D} \quad (4.13)$$

This may be compared to (4.5) for the exponential DSS. A better comparison is achieved if we plot borderline stability graphs for both types of DSS, for different sets of parameters σ, i, T_D . This is done in figure 4.9, with i on the x axis, T_D on the y axis, for four different values of σ . The graphs for the exponential DSS case are solid, while the annuity case graphs are dash-dotted. From the graphs we observe, as expected, that *cet.par.*, high interest rates or long loan durations give instability (i.e. debt growth, financial accumulation), for both types of DSS. And as already pointed out, an FRC closer to 1 gives debt growth, *cet. par.* We observe that the graphs for both types of DSS lie fairly close and have similar shapes (all graphs are hyperbolae). This gives support to the notion that the exponential DSS may be used for studying debt growth dynamics instead of the less algebraically tractable annuity DSS.

In the figure, loan duration T_{D2} for the annuity DSS has been adjusted in relation to T_{D1} for the exponential case, following the argument in conjunction with figure 4.3. In the figure, the T_D on the y axis = T_{D1} . By

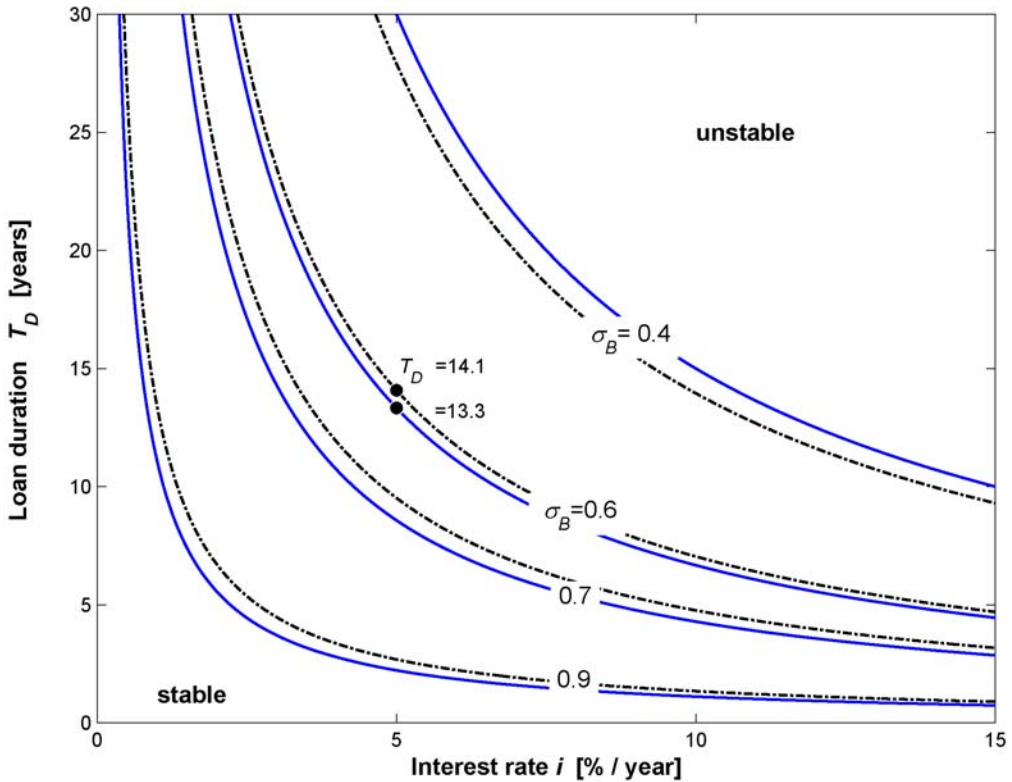


Figure 4.9: Regions of debt growth ("instability") for values of σ_B, i, T_D

experimenting it was established that $T_{D2} = 1.6T_{D1}$, not $T_{D2} = 2T_{D1}$ as suggested in figure 4.3, gave the best coincidence for the graphs over a reasonable range of values of σ . This adjustment does not, however, invalidate the use of the exponential DSS instead of annuity DSS, since the stability properties of both are so similar.

As an example of how stability information may be extracted from the figure, it is seen that at an interest rate of 5%/y and $\sigma = 0.6$, a loan duration $T_{D1} = T_D > 13.3$ will give accumulation when the DSS is exponential, and loan duration $T_{D2} > 1.6T_D = 1.6 \cdot 14.1 = 22.6$ gives accumulation for the annuity DSS case.

4.3.3 Firms with no income during a start-up period

The model of the firm sector presented here assumes that money arrives to firms in the form of demand for consumption and investment goods. This presupposes that the firm sector delivers a corresponding flow of products in the opposite direction. How then account for a not untypical situation where a firm receives a loan, but for a fair amount of time will not have any further monetary inflow since it has no products to deliver during its build-up phase?

Essentially, the solution is to modify the time profile of debt service, i.e. the impulse response of the debt service subsystem (DSS). If a new loan is extended at $t = 0$, the impulse response of the DSS is now set to zero for an initial period T (perhaps in the order of a year). The firm is exempt from debt service in this period. After $t = T$, debt service starts and follows the same profile(s) as already discussed, but after the original loan has first been amplified by a factor e^{iT} since compound interest must be added before debt service starts. Conditions for accumulation with this modified debt service profile changes somewhat, but the changes are not important for the analysis and are quite simple to derive. We will modify the exponential debt service scheme in eq. (4.1) so that it has the above properties (we could have done the same with the annuity scheme, but it does not make any significant difference for our analysis). The modified transfer function is

$$h_{D2}(s) = e^{iT} e^{-Ts} \frac{1 + iT_D}{1 + T_D s} \quad (4.14)$$

The term e^{iT} accounts for amplifying the debt, and e^{-Ts} accounts for the time delay before debt service starts. Since $h_{D2}(s)$ is irrational due to the term e^{-Ts} , we use the Nyquist criterion to check stability. Following a similar argument as that leading to (4.11), we now get

$$h_0(s) = -\sigma e^{iT} e^{-Ts} \frac{1 + iT_D}{1 + T_D s} \quad (4.15)$$

Again we may confine ourselves to considering (4.15) for $s = j\omega$ with $\omega = 0$. We have

$$h_0(j0) = \left[-\sigma e^{iT} e^{-j\omega T} \frac{1 + iT_D}{1 + T_D j\omega} \right]_{\omega=0} = (\text{real}) = -\sigma e^{iT} (1 + iT_D) \quad (4.16)$$

The system is unstable (i.e. accumulation occurs) for $-\sigma e^{iT} (1 + iT_D) < -1$. This corresponds to conditions for accumulation resembling those in

(4.5):

$$\sigma e^{iT} > \frac{1}{(1 + iT_D)}, \text{ or equivalently:} \quad (4.17a)$$

$$i\sigma e^{iT} > \frac{1}{T_D}(1 - \sigma e^{iT}), \text{ or} \quad (4.17b)$$

$$iT_D > \frac{(1 - \sigma e^{iT})}{\sigma e^{iT}} \quad (4.17c)$$

As expected, relieving firms of debt service for an initial period with the loan growing correspondingly, moves the system somewhat closer to the instability border for the same set of the three parameters interest, loan duration and banks' financial re-investment coefficient. Comparing (4.17) to (4.5), we see that stability-wise, a model with debt relief in an initial period, is equivalent to amplifying the FRC to $\sigma = \sigma e^{iT}$ in the original model (4.1).

With debt service relief in an initial period and the extreme special case $\sigma e^{iT} > 1 \iff \sigma > e^{-iT}$, conditions (4.17) tell us that accumulation will always occur.

4.3.4 Can capitalists earn any profits when money is debt-created?

There is a view among many Post Keynesians and Circuitists – see for instance Rochon (2005) – that if investment is debt-financed, the borrowing capitalists cannot receive any profits from their firms after servicing debt. At best, seen from the capitalists' position, it is a zero-sum game in the sense that non-financial capitalists that receive profits can only achieve this at the expense of other non-financial capitalists who run corresponding deficits.

We may discuss this using the model in figure 4.1. Simply put, if

$$F_i > F_D, \quad (4.18)$$

$F_o = \frac{M}{T}$ will be > 0 , and profits and wages are possible. The capitalists' share π of output after debt service depends on the balance of power between owners and employees. The profit flow is

$$\Pi = \pi \frac{M}{T} \quad (4.19)$$

The size of the profit share π is peripheral to this discussion, since as long as there is *some* money left in the real economy after debt service, a profit

flow > 0 is feasible. Profits (and wages) > 0 will be possible indefinitely if

$$\sigma < \frac{1}{(1 + iT_D)} \quad (4.20)$$

cf. condition (4.5).

In the model discussed here, both M and F_D result from the accumulation process and are outside debtors’ control. But T is a behavioural variable for this aggregate. Owners can achieve $F_o = \frac{M}{T} > F_D$ by reducing T , cf. corollary 2.6. In a setting with slow instability (accumulation), this is not possible indefinitely – there is some practical lower limit for T (correspondingly: some upper limit for real economy transaction frequency $v = 1/T$).

But the issue was here only to make the point that, in principle, capitalists in the aggregate may perfectly well enjoy a profit flow after servicing debt. This position is also supported by (Keen; 2010), making the point that this is a question – also touched upon earlier in this thesis – of avoiding the confusion of stocks (the initial loan) and flows (wages and profits). The reader is also referred to the simple model presented earlier, in subsection 2.7.5.

4.3.5 An intuitive, brief description

At this stage we will give an intuitive, non-mathematical description of the process that may lead to debt-induced crisis. See figure 4.10. The positive feedback from debt service to new loans is indicated with plus signs. Wages, dividends and expenses paid by the financial sector may then be considered a beneficial "leakage" back to the real (the non-financial) economy that weakens the accumulation process. In this sketch all households (also those working in finance) are located in the lower gray circle. The upper white circle does not contain any agents, but indicates a bundle of flows and their role.

Note that the debt service money flowing to lenders is in its entirety returned to the real economy. Abstracting from the effects of some lag in the financial sector this means that money will not disappear from the real economy, all of it will be cycled back. Due to accumulation however, it will to an increasing degree come back with strings attached – appearing as added debt. So we have a growth of the positive feedback flows to and back from the financial sector that leaves the aggregate of flows (Y) within the real economy behind.

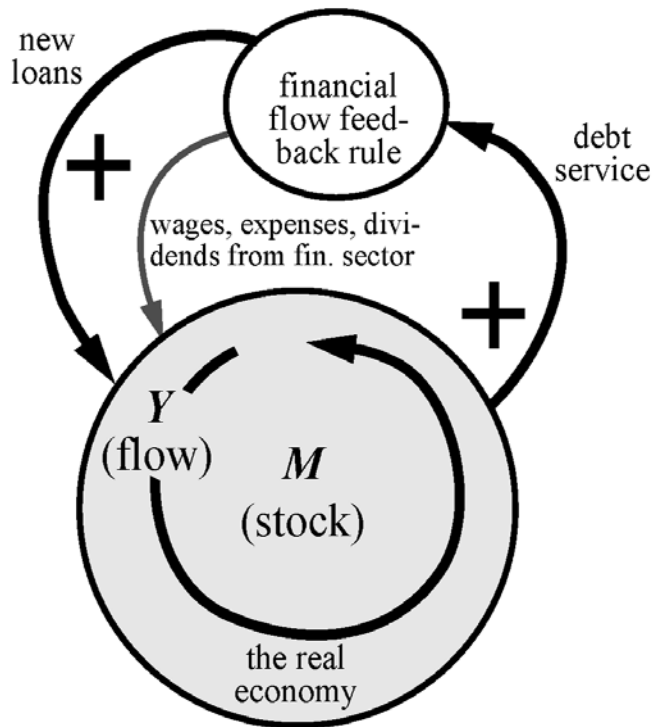


Figure 4.10: Debt polarisation

Debt service is a non-discretionary flow while the Y flows are discretionary, at least within some bounds. *When non-discretionary flows become dominating, the economy as a whole becomes more fragile.* This is the basic premise of the crisis mechanism we will explore later on.

4.3.6 Real growth does not save the system

A "law" of macroeconomics formulated by Thomas Piketty (Piketty and Goldhammer; 1995) is that as long as the real growth rate " g " is larger than the real rate of return on capital, the economy cannot end up in crisis (ignoring here the issue of ecological consequences of long-term persistent real growth). We will now discuss this briefly: The best-case situation is then when the real growth rate is high. Assume that that this unfolds in parallel with debt-asset polarisation. Inflation will be low due to high real growth. But the economy will still experience increased fragility due to

the relative growth of the non-discretionary (debt service) flows – high real growth does not make a fundamental difference for this.

One mechanism that might ameliorate the situation, is money growth keeping up with debt growth. Then a high real growth rate may at least stave off the inflation that is due to this money growth. The type of lending described in this chapter however, does not incorporate money growth. With such a model, ending up in debt-induced crisis is unavoidable when condition (4.5) is fulfilled, regardless of Piketty’s g .

A more sophisticated model incorporating money growth – where outcomes may be better – will be presented and examined in the next two chapters.

4.3.7 A measure of financial fragility

We return to figure 4.1, which in a mathematically more precise manner corresponds to figure 4.10 above. A measure of financial fragility θ may now be defined as debt service F_D divided by real economy output Y_o ($= F_o$ above):

$$\theta = F_D/Y_o \quad (4.21)$$

If this *debt service burden* increases above some percentage, a reasonable modeling assumption is that this will bring the economy into crisis: an aggregate of non-discretionary flow outgrows the aggregate of discretionary flows⁸.

But when crisis sets in, aggregate debt will increasingly be reduced through defaults, so that a feedback term should be introduced in the \dot{D} equation, like this:

$$\dot{D} = \dots - f(\theta)D \dots = \dots - \lambda D \dots \quad (4.22)$$

Here $\lambda = f(\theta)$ is some positive monotonic function, which will be called the *loss rate*. With such an added feedback, the system becomes nonlinear. In a crisis situation the loss rate could possibly have an ameliorating effect, countering the effect of a large θ . We will explore this through simulations in chapter 6.

We may – as an initial linear approximation – add a constant loss rate to (4.7), which then becomes

$$\dot{D} = (i\sigma - \lambda)D \quad (4.23)$$

⁸Income distribution is abstracted from in this thesis. Whether too much debt is to firms or households (and even the government), or a mix – it is considered to have a similar impact. This is as opposed to the Steve Keen’s Goodwin cycle models, where firm owners are assumed to be the sole borrowers. More on this model in section 6.3.

This equation shall be compared with an important equation (5.8) that is derived in the next chapter.

4.3.8 Summing up

In this chapter we have discussed lending dynamics for a system with constant money stock, where the lending sector may accumulate or not – based on the share of received debt service flow is re-lent, the interest rate, and the duration of loans. The outcome is either (slowly) exploding debt, or decaying debt. (The in-between "equilibrium" may occur only for a specific set of parameter values.)

Empirical data show that macroeconomies world wide have not been in an "indebtedness equilibrium" or a path with shrinking debt. Indeed, aggregate debt has persistently grown, and steeper than GDP. See figure 4.11. These "slow explosion" dynamics have largely been overlooked since the yearly growth rate has been low, and – more importantly – mainstream macroeconomics have mostly abstracted from debt (burden) dynamics.

The "moneylender" model in this chapter however, is not sufficient. *Money actually grows.* And the loan growth behind money growth is decided in a more complex manner, mainly by capital adequacy rules, which is treated in the next chapter. Later we will also account for the effects of the general mood among economic agents (both potential borrowers and lenders) which again is a function of other endogenous variables.

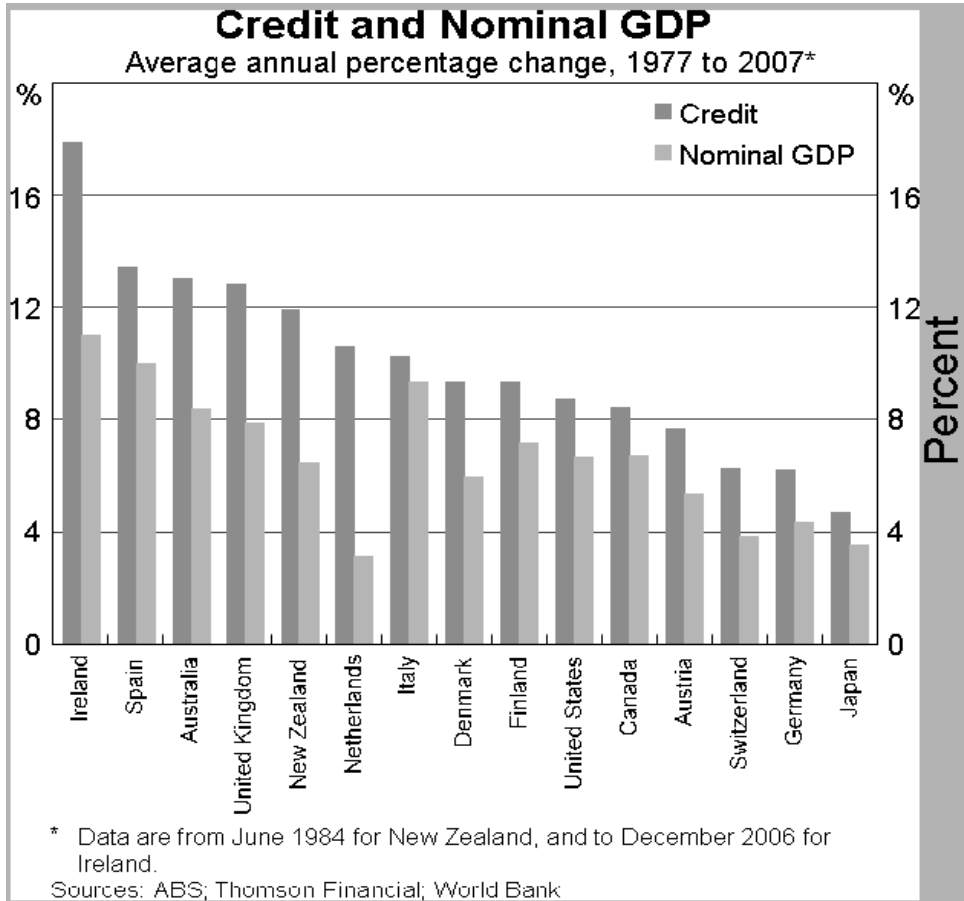


Figure 4.11: Debt outruns GDP in OECD countries (*courtesy: Reserve Bank of Australia*)

Chapter 5

A modern financial system and dynamics

Today's international regulation regimes for banks have been moving away from minimum reserve requirements and instead to demand that a bank's claims on others must exceed others' claims on the bank by some reasonable margin. More precisely expressed, banks are required to stay above some given lower bound for their capital/asset ratio (capital adequacy), also incorporating some risk weighing of different types of assets. This is a regulatory regime that stems from the Bank for International Settlements (BIS), and which is implemented nationally in most developed countries. We will in between call this the "BIS regime".

The only (acknowledged) rationale for this banking regulation regime – like the older one of minimum reserve requirements – is robustness against insolvency and bank runs. Here we will examine a side effect of a minimum capital/asset ratio regime. This chapter shows that it allows endogenous credit money¹ growth. It turns out that the maximum growth rate is inversely proportional to the required minimum capital/asset ratio, and that staying at this limit allows the steepest profit growth (abstracting from losses).

The model to be discussed is very simple, as indicated by the assumptions made below. Hopefully, it still embodies the properties needed for the analysis to be of value.

The bank's liabilities are its deposits. We assume no physical currency

¹We distinguish between credit money which is created through bank lending, and Central Bank money (reserves, HPM, base money). Credit money appears in the form of the public's demand deposits with commercial banks.

(bills and coins) in circulation, so that money stock is simply the aggregate of deposits. We initially also abstract from the presence of a government, a Central Bank and bank reserves. Except for government bonds, we also abstract from BIS rules for risk weighting of different types of assets. Until further notice we assume that all lending is done by commercial (licensed) banks – as opposed to non-banks (NBFIs) which we will introduce in the next chapter.

Later in this chapter we introduce a Central Bank and a government.

5.1 A generic bank model without a Central Bank

All monetary entities are as always in this thesis in nominal terms. We define the following variables and parameters for our bank:

$D(t), M(t)$ = assets, liabilities [\$]. M = money stock, as stated above.

κ_0 = the required minimum capital/asset ratio []

i_D = interest rate on assets (= loans = debt) [1/y].

i_M = interest rate on liabilities (= deposits = money) [1/y]; $i_M < i_D$

i = “equivalent net interest rate” (explained below) [1/y]

r = loan repayment rate [1/y]. r is defined such that the loan repayment flow is proportional to the loan, as discussed in the previous chapter.

λ = loss rate [1/y]; a flow $\lambda D(t)$ is written off due to borrowers defaulting on their loans

σ = share of net interest income that is left for banks after they have paid their expenses including wages []; $0 < \sigma < 1$. In "expenses" we also include investment-type outlays to increase the productivity of the banks. Banks are assumed to pay no taxes at this stage.

F_{nl} = flow of new loans [\$/y]

We assume that the general mood among lenders and borrowers is not pessimistic (when both sides or one side hold back). In that sense the models we are discussing here are only valid in a phase before crisis sets in. Then banks lend as much as they are allowed to, i.e. they (manage to) stay at the lower limit κ_0 . We have

$$\kappa_0 = \frac{D - M}{D}, \text{ or } D - M = \kappa_0 D, \text{ or } M = (1 - \kappa_0)D \quad (5.1)$$

The differential equation for asset change is

$$\dot{D} = F_{nl} - \lambda D - rD \quad (5.2)$$

The differential equation for liability change is

$$\dot{M} = F_{nl} - rD - \sigma(i_D D - i_M M) \quad (5.3)$$

Note that net bank income $\sigma(i_D D - i_M M)$ appears with a minus sign in \dot{M} , not with a plus sign in \dot{D} : net income to the aggregate of banks appears in the form of reduced bank liabilities (= customers' deposits). Using the rightmost equation in (5.1), the last term in (5.3) becomes

$$-\sigma [i_D - i_M(1 - \kappa_0)] D = -\sigma i D, \text{ where } i = i_D - i_M(1 - \kappa_0) \quad (5.4)$$

Here i may be termed an “equivalent net interest rate”. In i we now also include all types of fees on borrowers and depositors. These fees are assumed proportional to D and M , and may therefore be considered to represent an extra interest-like income for the bank.

We substitute (5.4) in (5.3), and substitute for M with the the rightmost variant of (5.1). This gives

$$\dot{D}(1 - \kappa_0) = F_{nl} - rD - \sigma i D \quad (5.5)$$

We subtract (5.5) from (5.2) and divide the result by κ_0 on both sides. This gives

$$\dot{D}(t) = \frac{\sigma i - \lambda}{\kappa_0} D(t), \quad (5.6)$$

which has the solution

$$D(t) = D_0 e^{gt}, \quad (5.7)$$

where we have introduced the aggregate assets growth rate

$$g = \frac{\sigma i - \lambda}{\kappa_0}, \quad (5.8)$$

and D_0 is the value of the bank's assets at $t = 0$.

We note that g increases with the equivalent net interest rate and the bank's profit share of income σ , which is not surprising. Note also from

(5.8) that g is very sensitive to λ ; we need $\sigma i > \lambda$ for assets to grow at all. A more interesting result is that

the growth rate is inversely proportional (5.9)
to the minimum capital/asset ratio

This growth rate also applies to the money stock, since we have $M = (1 - \kappa_0)D$ from (5.1) and may differentiate this on both sides:

$$\dot{M}(t) = (1 - \kappa_0)gD_0e^{gt} = \frac{(1 - \kappa_0)\sigma i - (1 - \kappa_0)\lambda}{\kappa_0} D_0e^{gt} \quad (5.10)$$

We observe that endogenous credit money growth will occur for $\kappa_0 < 1$. This is (as far as this author knows) a non-recognised side effect of a Basel-type regime. If $\kappa_0 = 1$ we have no money growth. Banks only lend out already existing money.

Using (5.3), (5.4), (5.7) and (5.8), the bank's net lending flow $F_{nl} - rD$ is

$$\begin{aligned} F_{nl} - rD &= \dot{M} + \sigma(i_D D - i_M M) & (5.11) \\ &= (1 - \kappa_0)gD + \sigma i D = \frac{\sigma i - (1 - \kappa_0)\lambda}{\kappa_0} D_0e^{gt}, \end{aligned}$$

which we will return to further below. Comparing (5.10) and (5.11), we note that the net lending flow is somewhat larger than the net money creation flow \dot{M} , which is reasonable since the bank also lends its own profit flow, and this is not accompanied by net creation of money.

The bank's profit flow is the difference between (5.11) and (5.10):

$$\sigma i D(t) = \sigma i D_0 e^{gt}, \quad (5.12)$$

That this flow grows steeper the lower the capital/asset ratio is, explains banks' wish – at least in optimistic times – to operate at the limit κ_0 .

5.1.1 A realistic and interesting result?

See figure 5.1². The thick graph gives an aggregate indicator of M1 growth in the OECD countries. (The other graphs are for individual countries – we

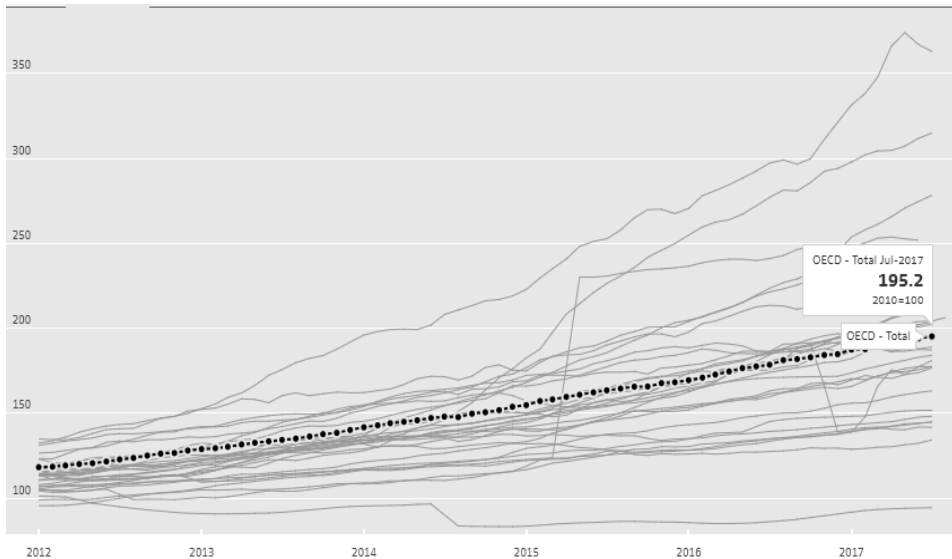


Figure 5.1: Aggregate indicator for M1 growth in OECD countries

will ignore them here). The growth rate extracted from the figure is 8.6%. For comparison, we try a set of numerical values to check out g from (5.8) above:

$$i_D = 0.055, i_M = 0.03, \sigma = 0.28, \lambda = 0.002 \text{ and } \kappa_0 = 0.08. \quad (5.13)$$

This gives a growth rate $g = 7\%$ per year. The point of this exercise is simply to check out whether the theoretical g from (5.8) calculated with some reasonable parameter values, is of the same order as indicated by the (for different countries admittedly *quite spread out*) empirical data for M1 growth. This is the case. Compare (5.8) to (4.23) in the previous chapter. We observe how a BIS regime allows dramatically steeper growth (more than ten times with $\kappa_0 = 8\%$) than what we get in the "moneylender" scenario in the last chapter, which would give much too low growth rates compared to empirical data – in the fraction of a percent range.

(There is no taxation of banks here. But $\sigma = 0.28$ is fairly low, so even if there is no government in the model, the effect of taxation may be thought of as incorporated.)

²(OECD; 2017)

We have a ballpark result that is reasonable. But is it interesting? To discuss this, we will consider a recent debate between some well-known economists. But first, let us refer to what is taught in many macroeconomic textbooks, and which is a common belief among economists, financial commentators and of course the public. Here, from a much-used economics textbook:

When banks loan out some of their deposits, they increase the quantity of money in the economy (Mankiw; 2008), p.355.

Yes, money supply increases when banks lend, as already argued. *But banks do not lend out of people's deposits.* Deposit accounts are a *liability* of the bank, as indicated by the role of M in (5.1) above – not something for a bank to lend.

That banks lend out of other customers' deposits – the *loanable funds* theory – is increasingly contested among economists and bankers. Paul Krugman, in a debate with Steve Keen, supports the loanable funds theory:

Keen then goes on to assert that lending is, by definition (at least as I understand it), an addition to aggregate demand. I guess I don't get that at all. If I decide to cut back on my spending and stash the funds in a bank, which lends them out to someone else, this doesn't have to represent a net increase in demand. Yes, in some (many) cases lending is associated with higher demand, because resources are being transferred to people with a higher propensity to spend; but Keen seems to be saying something else, and I'm not sure what. I think it has something to do with the notion that creating money = creating demand, but again that isn't right in any model I understand (Krugman; 2012)

Steve Keen replies to Krugman:

From the neoclassical vision of saving as modelled by Krugman (after inserting an implicit banking sector into Krugman's bank-less model), lending makes no difference to the level of aggregate demand (unless the impatient agent has a markedly higher propensity to spend) because lending does not change the amount of money in circulation—it only alters its distribution by reducing the amount in Patient's account and increasing

the amount in Impatient's. The banking sector's assets are ignored because the bank is treated as a "mere intermediary" that facilitates the loan between depositors (and maybe charges a fee for the service) but otherwise does nothing.

Real world lending is not a transfer of money from one depositor's account to another's, but a contract between a bank and a borrower in which the bank credits the borrower's account (thus increasing the bank's liabilities). In return for the borrower agreeing to be in debt to the bank for the same amount (thus increasing the bank's assets). This increases the aggregate amount of money in circulation, increasing aggregate demand in the process—and predominantly finances investment or speculation rather than consumption (Keen; 2014).

Lately, also Central Banks – traditionally quite reticent in airing "controversial" views in academic debates – have voiced opinions on this issue. Here is the Deputy Governor of the Norwegian Central Bank.

So how do banks create money? The answer to that question comes as quite a surprise to most people.

When you borrow from a bank, the bank credits your bank account. The deposit – the money – is created by the bank the moment it issues the loan. The bank does not transfer the money from someone else's bank account or from a vault full of money. The money lent to you by the bank has been created by the bank itself – out of nothing: fiat – let it become.

The money created by the bank does not disappear when it leaves your account. If you use it to make a payment, it is just transferred to the recipient's account. The money is only removed from circulation when someone uses their deposits to repay a bank, as when we make a loan repayment (Nicolaysen; 2017).

Also fairly recently, the German Bundesbank:

... banks can create book money just by making an accounting entry: according to the Bundesbank's economists, "this refutes a popular misconception that banks act simply as intermediaries at the time of lending – ie that banks can only grant credit using funds placed with them previously as deposits by

other customers". By the same token, excess Central Bank reserves are not a necessary precondition for a bank to grant credit (and thus create money) (Bundesbank; 2017).

And the Bank of England, which in fact was early (2014) among Central Banks in arguing this view:

The vast majority of money held by the public takes the form of bank deposits. But where the stock of bank deposits comes from is often misunderstood. One common misconception is that banks act simply as intermediaries, lending out the deposits that savers place with them. In this view deposits are typically ‘created’ by the saving decisions of households, and banks then ‘lend out’ those existing deposits to borrowers, for example to companies looking to finance investment or individuals wanting to purchase houses

In fact, when households choose to save more money in bank accounts, those deposits come simply at the expense of deposits that would have otherwise gone to companies in payment for goods and services. Saving does not by itself increase the deposits or ‘funds available’ for banks to lend. Indeed, viewing banks simply as intermediaries ignores the fact that, in reality in the modern economy, commercial banks are the creators of deposit money. This article explains how, rather than banks lending out deposits that are placed with them, the act of lending creates deposits — the reverse of the sequence typically described in textbooks (McLeay, Radia and Thomas; 2014).

This thesis supports and adds to this analysis, by demonstrating not only that money is created *ex nihilo* when banks lend. It also

1. derives the mathematical connection between the minimum capital/asset ratio requirement and the maximum endogenous growth rate of money (and profit, debt), and
2. that this theoretical growth rate is of the same magnitude as what is actually occurring, and suggesting that banks – at least in good times when there are enough willing borrowers – manage to stay close to the minimum capital/asset ratio limit.

5.2 Including a Central Bank

We now introduce a Central Bank (CB) and banks' reserves (banks' deposits with the CB). It is assumed that banks' reserves fluctuate with government spending and taxation (also introduced), and grow due to interest paid for these deposits.

Any CB where the country in question has its own national currency (as opposed to for instance the eurozone), is in the final instance an arm of the government. Hence, a government's "debt" that builds up with its CB through deficit spending in excess of the proceeds from selling bonds, is only an accounting and legal convention. In line with this, the government is in this thesis considered to be able to spend at its own discretion (and thus net create money) by debiting its account at the CB. (A possible economic impact of this type of net HPM creation is of course inflation, but that is no more an issue than the possible inflationary effect of banks' exponential net money creation, established in the previous section and also in the following.)

5.2.1 The bank model with a Central Bank and taxation

We distinguish between risk weight of reserves (zero) and all other assets in the Basel rule (these are for simplicity assigned a 100 % risk weight). We now define some additional parameters:

$R(t)$ = reserves = the bank's deposit with the CB = high-powered money (HPM) [\$]. We assume that $R > 0$. The bank's total financial assets are now $D + R$, where D = loans as before.

$\gamma(t)$ = government net spending (= deficit) flow. It may be negative, corresponding to a surplus budget. γ is an exogenous fiscal control variable for the system [\$/y].

β = tax rate for banks, $0 < \beta < 1$. All other taxation in the model is incorporated in $\gamma(t)$ [].

$\bar{\sigma}$ = profit share for banks after tax, $\bar{\sigma} = \sigma(1 - \beta)$ [].

i_R = interest rate on HPM deposits to banks from the CB. This is an exogenous monetary control variable for the system [1/y].

\bar{i}_R = after-tax interest rate on HPM to banks from the CB, $\bar{i}_R = i_R(1 - \beta)$ [1/y].

Remembering the Basel rule that risk weights shall only apply in the denominator, we get

$$\kappa_0 = \frac{D + R - M}{D + 0 \cdot R}, \text{ or } D + R - M = \kappa_0 D, \text{ or } M = (1 - \kappa_0)D + R \quad (5.14)$$

The differential equation for non-reserve asset change is

$$\dot{D} = F_{nl} - \lambda D - rD \quad (5.15)$$

The differential equation for change in the bank's reserves is

$$\begin{aligned} \dot{R} &= i_R(1 - \beta)R - \beta[\sigma(i_D D - i_M M) - \lambda D] + \gamma \\ &= \bar{i}_R R - \beta\sigma(i_D D - i_M M) + \beta\lambda D + \gamma \end{aligned} \quad (5.16)$$

The bank pays taxes by drawing on its reserves R . Taxes are assumed to be paid both on interest income from R , and from the profits extracted from its banking activities after losses λD are subtracted.

The differential equation for liability (= deposit money) change now becomes

$$\dot{M} = F_{nl} - rD - \sigma(i_D D - i_M M) + \gamma, \quad (5.17)$$

where the second last term in (5.17) may, using (5.14), be written as

$$-\sigma[i_D - i_M(1 - \kappa_0)]D + \sigma i_M R = -\sigma i D + \sigma i_M R, \quad (5.18)$$

where $i = i_D - i_M(1 - \kappa_0)$ as before. Using (5.18), (5.17) becomes

$$\dot{M} = F_{nl} - rD - \sigma i D + \sigma i_M R + \gamma, \quad (5.19)$$

We substitute for \dot{M} in (5.19), using a differentiation of the rightmost variant of (5.14), and also substitute (5.16) for \dot{R} . This gives

$$\dot{D}(1 - \kappa_0) + \bar{i}_R R + \gamma - \beta\sigma(i_D D - i_M R) + \beta\lambda D \quad (5.20)$$

$$= F_{nl} - rD - \sigma i D + \sigma i_M R + \gamma, \quad (5.21)$$

where γ cancels out on both sides, $\bar{i}_R R$ and $\beta\lambda D$ may be moved to the right side, and we use $\bar{\sigma} = \sigma(1 - \beta)$:

$$\dot{D}(1 - \kappa_0) = F_{nl} - rD - \bar{\sigma}iD - \beta\lambda D + \bar{\sigma}i_M R - \bar{i}_R R \quad (5.22)$$

We subtract (5.22) from (5.15) and divide the result by κ_0 on both sides. This gives

$$\dot{D}(t) = gD(t) + \frac{\bar{i}_R - \bar{\sigma}i_M}{\kappa_0} R(t), \text{ where } g = \frac{(\sigma i - \lambda)(1 - \beta)}{\kappa_0} \quad (5.23)$$

Compare this to equation (5.6). The growth equation has a similar structure, but it is now also influenced by the variable R , whose growth is decided by the two control variables/parameters γ and \bar{i}_R in (5.16). The growth rate g will also be lower the higher the tax rate, since we have $\bar{\sigma} = \sigma(1 - \beta)$.

But D and M will still grow exponentially. For the system to uphold the balance between monetary aggregates, R must grow at the same rate. If R is depleted, banks will increasingly lack reserves for their transactions with each other, to the government and the public (for notes and coins). Persistent government deficit spending, i.e. positive γ (what is usually talked about in admonitory terms as "printing money"), achieves net HPM creation. The same is the result of the CB buying government bonds to achieve an interest rate target.

The basic assumption in the above is that the government (or monetary authorities) allows a commercial bank and its potential borrowers to generate debt and money growth, the bank only being constrained by a BIS-type capital/asset-ratio requirement. And the Central Bank accomodates banks by ensuring reserve (R) growth.

But one might instead take the position that the monetary authorities should decide the rate of M growth. This can be achieved in a *100% reserve system*. In chapter 7 we will discuss this, and propose a solution variant which has the advantage of 100% reserves (money cannot be lost) but avoiding the main disadvantage (reduced flexibility and less easy availability of loans for investment).

5.2.2 A remark on government bonds

This thesis abstracts from the selling and buying of government bonds, and there is no corresponding financial market. The reason from this is as follows: the composition of reserves vs government bonds for a bank does not impact on the value of banks' capital/asset ratio, since government bonds

are weighted with zero in the denominator, like reserves. And when the government sells an x [\$] bond to the bank, the government adds x to its deficit spending flow γ just as if the government had deficit spent x by debiting its CB account.

By abstracting from government bonds (and corresponding debt) we are at this stage ignoring a component in the aggregate of non-discretionary flows in the economy – the part of taxes that are collected to service debt to bondholders. When we in the following examine mechanisms for debt crisis, we initially confine ourselves to non-discretionary flows due to private sector – not government – debt. But we will later argue, in subsection 6.1.1, that the introduced model *can easily incorporate government debt and debt service*.

Related to this: in subsection 7.2.1 we launch a proposal that largely removes the rationale for today's (dangerously powerful) government bond markets.

5.2.3 Money growth pari passu with debt – a good thing!

The assumption in this chapter has been that the government (or monetary authorities) allows the bank to decide debt and money (M) growth, only being constrained by a BIS-type capital/asset-ratio requirement (dependent also on borrowers' demand for loans).

But one might instead take the position that the monetary authorities – not banks – should decide the rate of M growth. Much of the criticism against banks – especially during the recent global financial crisis years – has been about banks creating money when extending loans. This has been contrasted with the (allegedly better) solution of money creation by the government: "spending money into existence" instead of "lending it into existence". This leads to a *100% reserve system*, where reserves mirror deposits 100%, i.e. $M = R$. It is the historically well-known proposal put forward by, among others, Irving Fisher during the Great Depression (Fisher; 1936), pp. 406-420. It has been persistently promoted to this day by individuals or groups that are more or less considered to belong to the economics "fringe", and has (in this author's opinion: undeservedly) not been considered worth serious discussion by the academic mainstream.

That said, we will now argue "heretically" (at least seen from a heterodox position) that bank money creation when lending is actually a *good* thing. Consider (5.14). Ignoring R (i.e. we then have the worst case for money growth in relation to debt growth), the relation is simply

$$M = (1 - \kappa_0)D, \quad (5.24)$$

so M and D grow at the same exponential rate. From earlier we have for output (2.44),

$$Y_o(t) = Mv, \text{ where } v \text{ is money velocity} \quad (5.25)$$

Assume now that money velocity v doesn't vary too much in non-crisis times, at least staying within the same magnitude³. The debt service burden introduced earlier in subsection 4.3.7, may be written as

$$\theta = \frac{(i+r)D}{Mv} = \frac{(i+r)D}{(1-\kappa_0)Dv} = \frac{(i+r)}{(1-\kappa_0)v}, \quad (5.26)$$

so θ does *not* increase exponentially – nominal Y grows (abstracting from shorter-term business cycles and external shocks) approximately *pari passu* with D . So how can one have D actually growing *faster* than Y as indicated by empirical data, shown in figure 4.11? This will be a topic in the next chapter.

5.3 "Bootstrapping" a money-creating bank

We will in this section try to explain how a newly established bank – "from scratch" – is able to create extra credit money when lending, at an exponentially growing rate. When a bank extends a loan, a corresponding increase occurs in the borrowing depositor's account. This means that the bank creates money "out of thin air" when lending. In principle, if a bank was unconstrained by regulation, it could create as much extra money (and debt) via this process as it wanted, as long as there were willing borrowers.

But banks are as already discussed constrained in their lending by the BIS rules on a minimum capital-asset ratio. The capital/asset ratio for our bank is chosen as the same simplified variant used earlier,

$$\kappa = \frac{D + R - M}{D}, \quad (5.27)$$

which must be $\geq \kappa_0$, the mandated minimum C/A ratio.

³M1 in for instance USA was not constant in pre-crisis times, but had a long-term velocity growth around 2% from 1960 to 2008, then diving sharply when crisis hit. The very simplified assumption here of constant velocity in pre-crisis times will therefore be discussed later, in subsection 6.2.7.

In our exercise we start at time zero with investors pooling their money to start a new bank. The bank has no outstanding loans and no deposit(or)s at $t = 0$. We assume that the bank's investors have supplied an initial capital R_0 at $t = 0$. The invested amount R_0 resides at the new bank's account at the Central Bank, and are thus reserves. For simplicity, we hold $R = R_0$ constant in the simulation interval. This is a fairly non-dramatic assumption with respect to M growth, and does not invalidate our argument.

Before proceeding, we list the variables and parameters that are used in this section. Some "benchmark" parameter values are chosen for this exercise, shown in {braces}:

$D(t)$ = the bank's outstanding loans [\$].

$M(t)$ = deposit money [\$].

$R(t)$ = the bank's reserves at the Central Bank [\$]

$K(t)$ = the bank's capital, $K = D + R - M$ [\$]

i_L = interest rate on loans given {0.055} [1/y].

i_M = interest rate on deposit money {0.03} [1/y],

λ = pre-crisis loss rate on loans {0.002} [1/y],

r = loan repayment rate {0.1} [1/y].

σ = share of net interest income that is left for the bank after payment of wages, dividends, taxes and other expenses {0.28} []; $0 < \sigma < 1$.
Note that we from now on, for simplicity, incorporate taxes in this σ . This is in contrast to the more complicated model used above, in subsection 5.2.1.

i : The bank's profit flow after all expenses, before losses,
 are $= \sigma(i_L D - i_M M) = \sigma(i_L D - i_M [1 - \kappa_0] D) = \sigma i D$

$F_{nl}(t)$ = flow of new loans from the bank [\$ / y]

κ_0 = required minimum capital-asset (C/A) ratio {0.08} []

$\kappa(t)$ = current actual C/A ratio [], $\kappa(t)$ must be $\geq \kappa_0$.

As a first step we assume – unrealistically – that our bank immediately lends an amount R_0/κ_0 precisely at $t = 0$. This gives the system initial values $D_0 = R_0/\kappa_0 = M_0$, which – following (5.27) – places $\kappa(0)$ exactly at κ_0 for $t = 0$, and then we will also have $\kappa(t) = \kappa_0$ for the whole time path. We choose a time interval of 50 years, which we will use for all later simulations. With these conditions, we may derive the debt and money creation path algebraically, by using equation (5.23). For simplicity we here assume no taxation and no interest on bank reserves R . We get a corresponding simplified version:

$$\dot{D}(t) = gD(t) - \frac{\sigma i_M}{\kappa_0} R_0, \text{ where } g = \frac{(\sigma i - \lambda)}{\kappa_0} \quad (5.28)$$

With the given initial values the solution is

$$D(t) = D_0 e^{gt} - \frac{\sigma i_M}{\kappa_0} R_0 \frac{1}{g} (e^{gt} - 1) = \frac{R_0}{\kappa_0} e^{gt} - \frac{\sigma i_M}{\kappa_0} R_0 \frac{1}{g} (e^{gt} - 1) \quad (5.29)$$

The resulting graph is shown in dashed line in figure 5.3 (we will comment on the other graphs further below).

A weakness of this algebraic approach is that we assumed that the bank was able to immediately lend the necessary amount to reach the allowed C/A ratio lower bound. This – and other improvements – can be achieved by formulating the bank's behaviour as a control system:

5.3.1 The bank as a control system

In this approach the mandated minimum capital/asset ratio κ_0 is a reference input, $\kappa(t)$ is a resulting output and the loan-issuing flow $F_{nl}(t)$ is the control action.

We choose a proportional plus integral (PI) controller, since integral action ensures that the control error (between desired minimum C/A ratio κ_0 and the actual ratio $\kappa(t)$) tends to zero. Even if the bank's officers obviously don't behave like a (non-human) controller in a technical system, we assume that they act in a similar way, to ensure as much lending as possible without being in breach of the minimum C/A requirement. This reasonably corresponds to the effect of PI control.

The "control system" is portrayed⁴ in figure 5.2. It contains the PI

⁴To indicate that an output is a function of an input, as opposed to a block that indicates that the output is just the input multiplied by the block content, a thick outline is used.

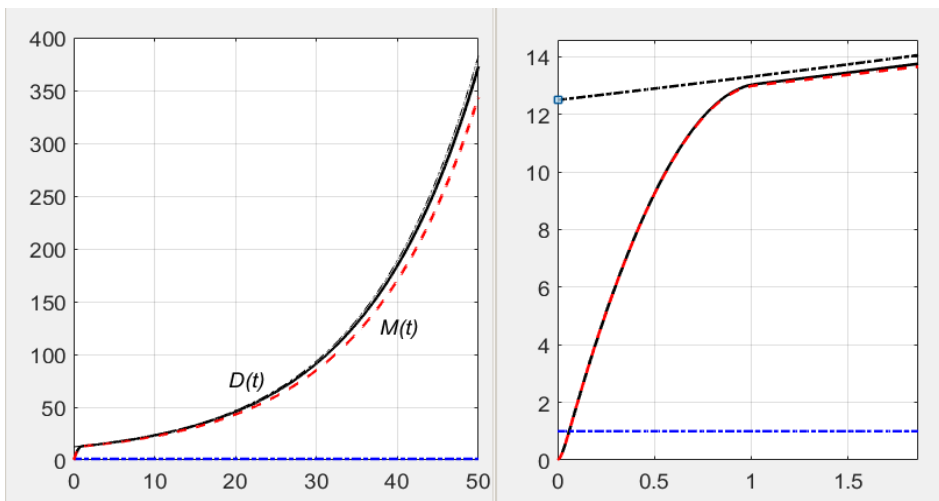


Figure 5.3: Debt and credit money growth without government and a CB. Right half is magnified, close to $t = 0$.

"out of thin air" along with debt D in the process. Parameter values are the same as in (5.13), and the simulation tends to the graph given by (5.29), as expected. The slight discrepancy between D given by (5.29) and the D graph resulting from ramping up via PI control, is due to the transient difference close to $t = 0$, as observed in the magnified right half of the figure. The horizontal dash-dotted line is the bank's reserve account R_0 , which for convenience is held constant $= 1$ in this exercise.

An objection to the above is that the bank might not be able to lend as much as it was allowed to by the capital adequacy rule. One might try to account for this by letting the loan flow be less than the $F_{nl}(t)$ output given by the PI controller, using a coefficient $0 < c_{nl}(t) < 1$ so that the modified loan flow is $c_{nl}(t)F_{nl}(t)$. This was tried, but didn't work. After some thought, the reason was obvious: decreasing the gain of the controller was compensated automatically by an increased control error, so $F_{nl}(t)$ was upheld at the allowed maximum level.

An alternative way out is to let the reference input $1/\kappa_0$ be multiplied by a similar factor, so that we now have a varying reference, $c_{nl}(t)/\kappa_0$. This has a reasonable interpretation. If a bank doesn't find or experience demand from enough customers, and/or potential borrowers are not considered creditworthy, it will reduce its lending ambitions, corresponding to $c_{nl}(t) < 1$.

This will be the case in economic downturns or crises, where the general mood turns more pessimistic: banks will hold back in their lending and potential borrowers will voluntarily choose to abstain. We will later on use κ_0 not as a constant, but as a variable dependent on other parts of the system.

The model introduced above will be used as a subsystem in the next chapter, where we will model and simulate dynamics leading to an an over-indebted economy, financial crisis and collapse.

Chapter 6

Debt build-up and crisis

We have by now looked at two types of banking systems: A "moneylender" system (non-bank financial institutions, aka non-banks) where debt may accumulate but money supply is constant. And the modern system with licensed commercial banks (only capital adequacy but no reserve requirements), where we have concluded that debt and money grow *pari passu*.

Before discussing crisis mechanisms, we will add to our commercial bank-based model from section 5.3, so that the model also accounts for the effects of non-bank lending. The commercial bank is now a submodel; the aggregate of commercial banks. What we call non-banks corresponds to the earlier used term NBFIs (non-bank financial institutions). Non-banks receive flows from a subgroup of the NFF (non-financial firm) capitalists who are owners of real economy firms, but who channel a share of their profits to financial investment with non-banks, not only to direct investments in the real economy.

This extended model – shown in figure 6.1 – will be used in the following exercise:

6.1 A "Minskyan" long-term crisis model

We continue with an economy where, at $t = 0$, we start "from scratch" with no credit money and a (from now on *aggregate*) of banks. This aggregate of banks (called "the Bank" with a capital B) has one type of asset, which is the aggregate of loans to households and real-economy firms; "the real economy". Since all banks are lumped together in an aggregate, lending between them cancels out.

This aggregate holds a constant small amount of reserves R_0 at the CB

(R_0 will still be held constant for simplicity – this does not impact much on the analysis to follow). The real economy starts borrowing from the Bank, which creates credit money *ex nihilo* at the exponential rate described in section 5.3. This enables the real economy to circulate money and achieve a corresponding exponentially increasing GDP at the same rate (as always in this thesis, in nominal terms). GDP is shared between worker households who are assumed to spend their whole income, and "capitalists" who spend some, and invest or lend the rest.

The time horizon is long, set to 50 years. At a some time into this period, there is a gradual change in the attitude to investment, so that capitalists increasingly prefer to invest financially (lending to the real economy), not invest profits directly into real-economy firms. They also increasingly prefer to channel their investment money to a financial asset sector (i.e. non-banks), not to the real economy. This will be expressed by a "financialisation coefficient" $0 < c_{FI}(t) < 1$, with effects described in the model further below. Through an initially small but growing c_{FI} , we first have a fifties/sixties-like real economy-dominated period (low c_{FI}), gradually changing into a later eighties/onwards "financialised" period (high c_{FI}). This coefficient is assumed to grow exponentially, so that there is a long initial period where it has negligible impact. The introduction here of such a coefficient is inspired by Hyman P. Minsky. In his works, he presents his *Financial Instability Hypothesis*. He briefly sums it up like this (Minsky; 1992):

The first theorem of the financial instability hypothesis is that the economy has financing regimes under which it is stable, and financing regimes in which it is unstable.

The second theorem of the financial instability hypothesis is that over periods of prolonged prosperity, the economy transits from financial relations that make for a stable system to financial relations that make for an unstable system.

The second theorem describes what in the debate today is called "financialisation".

c_{FI} is in our model given exponential – not linear – growth, because financialisation is a process of forgetting and of contagion. *Forgetting* the lessons from the last crisis, and *contagion* since agents (slowly but increasingly) transit to more reckless behaviour, following the herd.

Furthermore, there is a positive feedback regarding what we could call *re-lending pressure*: A persistently increasing income due to financial ac-

cumulation gives a persistently increasing incentive to throw loans after borrowers that are not creditworthy. This pressure adds to recklessness.

The new and extended model is shown in figure 6.1.

The additional variables and parameters in figure 6.1 are listed below. Parameter values for the Bank part of the model are the same as in section 5.3. Some parameters are assigned constant values in initial simulations. Their values are shown in {braces}. Other parameters vary from the outset, this is explained further below.

T_F = input-output time lag for the aggregate of non-financial firms (NFFs) {0.5} [y].

T_K = time lag for spending and real-economic investment for the aggregate of NFF capitalists, out of real-economic profits {0.3} [y].

T_{KF} = time lag for lending money from the aggregate of NFF capitalists [y].

T_W = time lag for the aggregate of (non-saving) workers/households, approx. 1 month = {0.08} [y].

$\dot{M}(t)$ = net flow of Bank-created money to the real economy [\$].

$F_{KL}(t)$ = flow of new loans to the real economy from NFF capitalists [\$/y].

$Y_d(t)$ = aggregate demand for non-bank products and services [\$/y].

$Y_o(t)$ = aggregate real economy output [\$/y].

$D_K(t)$ = NFF owners' (capitalists') financial assets = loans, bonds but not stocks [\$].

i_K = interest rate (net after taxes, fees and costs) on D_K {0.04} [1/y].

λ_K = non-crisis loss rate on D_K , always assumed 2.5 times the loss rate on bank loans {0.005} [1/y].

r_K = loan (bond) repayment rate {0.1} [1/y]. r_K is defined such that the loan repayment flow = $r_K D_K(t)$, proportional to the aggregate of NFF capitalist loans.

$\pi(t)$ = share of NFF output that capitalists receive []; $0 < \pi < 1$. The workers' share is then $1 - \pi$.

σ_K = share for real-economic investment as opposed to consumption share $1 - \sigma_K$, out of real economy profits not going to lending {0.25} []; $0 < \sigma_K < 1$.

σ_{KF} = share for capitalist lending as opposed to consumption share $1 - \sigma_{KF}$, this share out of their interest income plus repayment flow {0.75} []; $0 < \sigma_{KF} < 1$.

$c_{FI}(t)$ = "financialisation coefficient", the central time variable expressing degree of financialisation in the economy []; $0 < c_{FI} < 1$.

$\Pi_R(t)$ = profits from $Y_o(t)$ that capitalists receive [\$/y].

$\Pi_{RF}(t)$ = profits from $Y_o(t)$ that capitalists receive and allocate for lending [\$/y].

$\Pi_L(t)$ = net interest income plus repayment from lending that capitalists receive [\$/y].

We will build on the model in figure 6.1. First a caveat: We have already noted that the dimension of parameter space in the pre-crisis model is large, and we have by trial and error explored a few points in parameter space to get responses that seems reasonably realistic. This dimensionality problem grows further when we now endeavour to introduce additional interactions, functional relations and associated parameters.

To cut through this knot and not become bogged down in what could in itself constitute a separate dissertation, the following procedure has been chosen: Introduce a few new parameters and related interactions in stages, examine responses and adjust parameters to avoid absurd outcomes, and finalise with a parameter set that give acceptable responses.

Then add some further relations and repeat the procedure. All parameters are up for adjustment. Finalise the crisis model with a parameter set that give an intuitively "reasonable" outcome. Discuss. Constant parameters will in later simulation stages become variables, connected to other system variables. In the first simulation experiment only $c_{FI}(t)$ will vary, the other parameters are held constant.

We note that in this model – and in this thesis as a whole – there is no connection between the Bank and the non-banks. Non-banks here don't buy Bank-issued debt, and the Bank does not borrow from non-banks. This is a weak point, and something that should be pursued in further research. But it is not believed to have a decisive impact on the main insights and conclusions in this thesis.

6.1.1 Including a government and Central Bank

Before embarking on a series of simulations, we will argue that the model in figure 6.1 can easily incorporate a government and Central Bank, and that such an addition may be simplified to such a degree that the structure remains unchanged, with no significant loss of explanatory power:

At several points in the circulatory network in 6.1 the government extract tax flows, even if these are not shown. These government-related flows are split off and enter a time lag block (expressing part of government stock-flow dynamics), leave this block, and re-enter somewhere else in the same real-economy circulatory system. The extracted (split-off) flows – taken together – constitute an aggregate tax flow, and is returned via a collection of time lag government blocks as wages, pensions, subsidies, purchases – back to the real economy at different injection points.

All these flows are split off from and re-injected into the already defined circulatory flows, such as consumption and investment. By collecting all extraction points using rules for block diagram manipulation, and collecting all injection points using similar rules – and also using rules developed in this thesis for calculating aggregate time lags – we may achieve an equivalent government net aggregate flow entering at the main summation point indicated in front of the "firms" block in figure 6.1, without changing anything except (somewhat) the time constants T_F , T_W and T_K . This aggregated injection flow may be positive (the government runs a deficit) or negative (the government runs a surplus).

We may now account for government debt and related debt service by considering the government as embedded in the network of nodes (blocks) in the real economy. By this, total debt $D + D_K$ in the model *includes* government debt, and government debt service to the Bank and non-banks are correspondingly included in the debt service flows in the model.

What remains after these strong – but valid – simplifications, is simply a net government flow entering at one point in the circulatory system. We may then continue using the model in figure 6.1, with just an added government net flow at the main injection point. We return to this later in subsection 6.2.8

Running a surplus budget (note: "surplus" here defined as *net of borrowing and debt service*) corresponds to a negative flow, i.e. *money is extracted from circulation* ("blood is drained from the body").

Concerning the opposite case – a positive net flow – we have earlier stated in section 5.2 that the generic nation assumed in this thesis issues its own currency. Then the government may ensure a positive net government flow

by simply debiting its account at the Central Bank, "spending, not lending money into the economy". By this we follow the position of modern money theory, which rejects the dogma of "Central Bank independence". MMT stipulates that a government issuing its own currency can always spend the amount needed for full employment and full capacity utilisation by debiting its account at the Central Bank, i.e. creating money ex nihilo. This is politically controversial, but it will – based on the analysis in this thesis – be supported as a crucial control tool. We will come back to this in detail in chapter 7.

Summing up: The above simplified way of accounting for a government's role in the macro circulatory system, allows us to lump government and private sector debt together when we in the next section examine crisis dynamics.

6.2 Simulation experiments

We continue with a 50-year time horizon, as in section 5.3. We will do five simulations, and start with the simplest variant:

6.2.1 All time lags and loss rates constant, only varying c_{FI}

The central parameter is $c_{FI}(t)$, which will be used in functions that decide other parameters. Experiments with many values have been done, and resulted in this choice:

$$c_{FI}(t) = 0.05 \exp\left(\frac{\ln(0.55/0.05)}{50} t\right) \quad (6.1)$$

This is an exponential function giving $c_{FI}(0) = 0.05$ and $c_{FI}(50) = 0.55$. It expresses increasing financialisation in the economy over a long time horizon, and impacts on other parameters that are discussed below. c_{FI} will be given influences in several ways in the model. In this first run, the only impact of c_{FI} is to channel more of capitalists' real economy profit towards financial investment. We assume that real-economy profits received by the capitalist subgroup is increasingly channeled away from consumption/real-economy investment, and to lending. The real-economy profit flow that capitalists allocate to lending is

$$\Pi_{RF}(t) = c_{FI}(t) \pi(t) Y_o(t) \quad (6.2)$$

At $t = 50$ we have $c_{FI}(t) \pi(t) = 0.55 \cdot 0.2 = 0.11$. This is quite low. But note that the non-bank subgroup in addition to this also receives profits +

repayment Π_L from their lending. Note also that profit-related parameters express net flows *after taxes*.

All other parameters are kept constant, not impacted by $c_{FI}(t)$. The results are shown in figure 6.2.

Graph (c) shows how non-bank debt starts at zero, stays very low for many years and takes off after approximately 30 years. This is explained by the growth in c_{FI} , shown in figure 6.2(b). Total debt/GDP ratio is shown in graph (d). The increase to approx. 120% there is due solely to non-bank debt and not Bank debt, a point made early in section 5.3. But things proceed without crisis occurring, and the nominal output growth rate increases persistently towards 7%, corresponding to the endogenous growth rate we got from Bank loan and money growth in subsection 5.3.1. (Ignore the sharp vertical transient to the left in graph f); it is due to the system starting up from zero at $t = 0$).

We will now change the above idyllic scenario into a more realistic one (and other changes follow further below):

6.2.2 Increasing c_{FI} , also T_{KF} and π

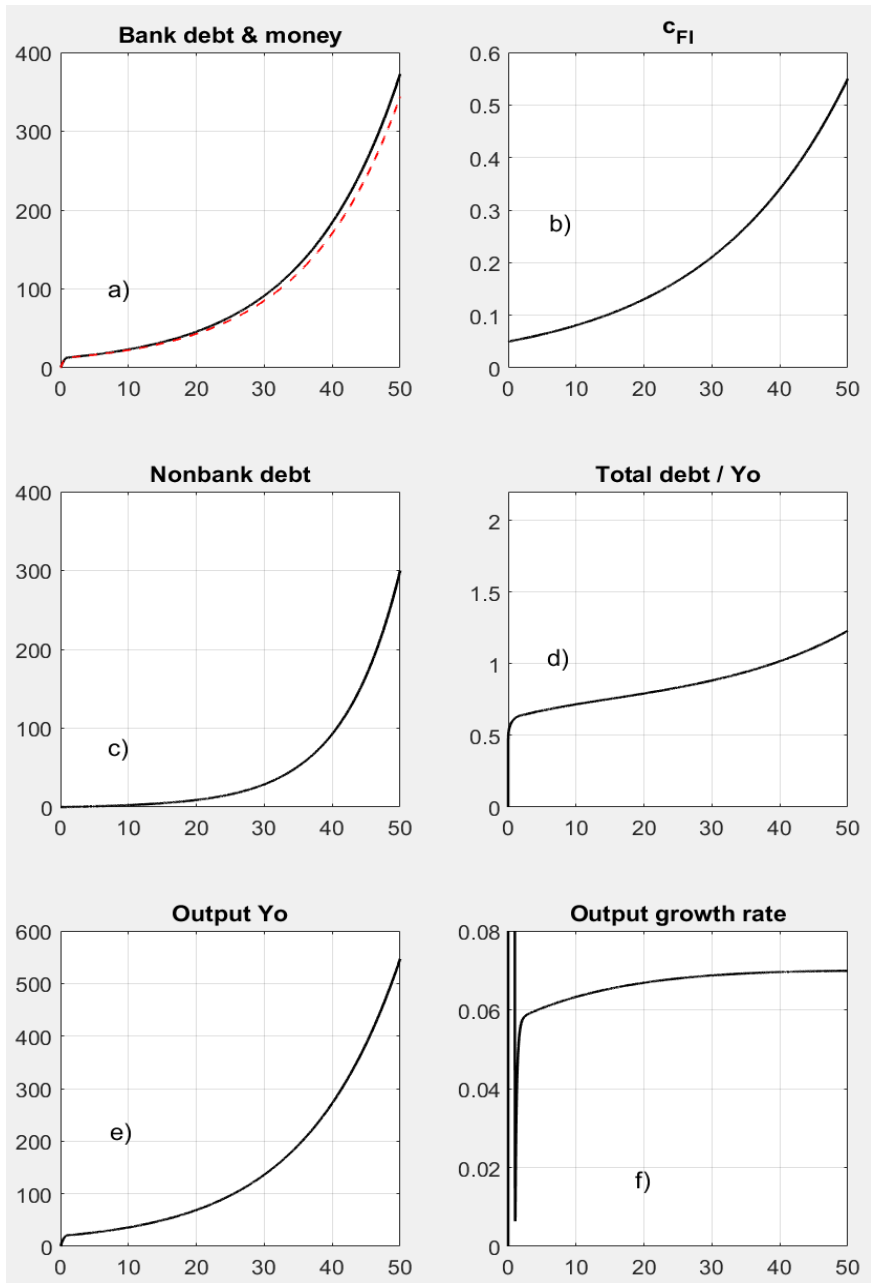
We now assume that increasing financialisation implies that more of the capitalists' income that is destined for lending is channeled to other capitalists, and not to the real economy. In our model this simply corresponds to a higher time lag T_{KF} . Money arriving to this subsystem spends more time there since it to larger degree circulates inside for intra-non-bank lending/borrowing and speculative trading in assets, before emerging as $F_{KL}(t)$. Based on this, we assume an increasing time lag

$$T_{KF} = 0.801 (1 + 5 c_{FI}(t)) T_{KF0} \quad (6.3)$$

where T_{KF0} is a constant, a fairly short time interval 0.3 [y].

(6.3) gives $T_{KF}(c_{FI}(0)) = 0.3$ [y] and $T_{KF}(c_{FI}(50)) = 0.9$ [y]. At $t = 50$, out of 3 arriving dollars to a non-bank capitalist, only one dollar is re-lent to the real economy, without first being channeled into the NFF capitalist subgroup. The "outside spending coefficient" ρ introduced in 2.8 falls with t , and is $1/3$ at $t = 50$.

Debt *within* the non-bank subgroup nets to zero. This adds to system fragility in the long run, but we will not go into this in this thesis. Ignoring it here gives a "best case scenario" regarding crisis dynamics. We restrict the treatment here to the increased time lag before money emerges from the NFF capitalist subgroup for new loans $F_{KL}(t)$ to the real economy.

Figure 6.2: Increasing c_{FI} only

An interesting point implied in this model is that "internal" non-bank lending and speculative trading within the NFF capitalist subgroup is mostly a zero-sum game for participants (one agent's win is another's loss), as opposed to lending to the real economy. One might object that "why then would non-bank capitalists lend to or invest financially with each other?" The answer is that if you believe that you are more clever than the other guy, you will do this even if your gain is dependent on others' loss; the financial market has a poker game psychology. Furthermore, observation of increasing flows into the subgroup attracts activity to it, a positive feedback mechanism.

This said, it should be added that speculation in real estate is not necessarily a zero-sum game, since it implies power for the owner to extract rent (a non-discretionary flow) from tenants who have no alternatives. This effect may be accounted for by the following: We now introduce a change in the power relationship between capitalists and workers, so that the profit share π of output Y_o increases with $c_{FI}(t)$. We choose

$$\pi(t) = (0.2 + 0.15 c_{FI}(t)) \quad (6.4)$$

This gives $\pi(0) = 0.21$ and $\pi(50) = 0.283$.

Simulation results are shown in figure 6.3.

If we compare figures 6.2 and 6.3, we observe that output growth culminates, see f), and the debt burden d) increases. Output growth culminates because money is drawn away from the real economy to circulate in the non-banks subsector, and because a larger share of output goes to capitalists. The increase in debt burden is due to two factors: larger non-bank debt c) and lower output e).

But still the scenario is too optimistic. Even with a much heavier debt burden, there is no increase in Bank loss rate λ and capitalist loss rate λ_K . There is also no holding back in non-discretionary spending by firms, workers and capitalists, even when increasing losses after some time also is recognised and felt among the general public. This is corrected in the third simulation:

6.2.3 Growing c_{FI} , profit share, time lags & loss rates

This is the full-on pessimistic but most realistic model. The system is now stressed by an increasing debt burden

$$\theta = \frac{D}{Y_o} \quad (6.5)$$

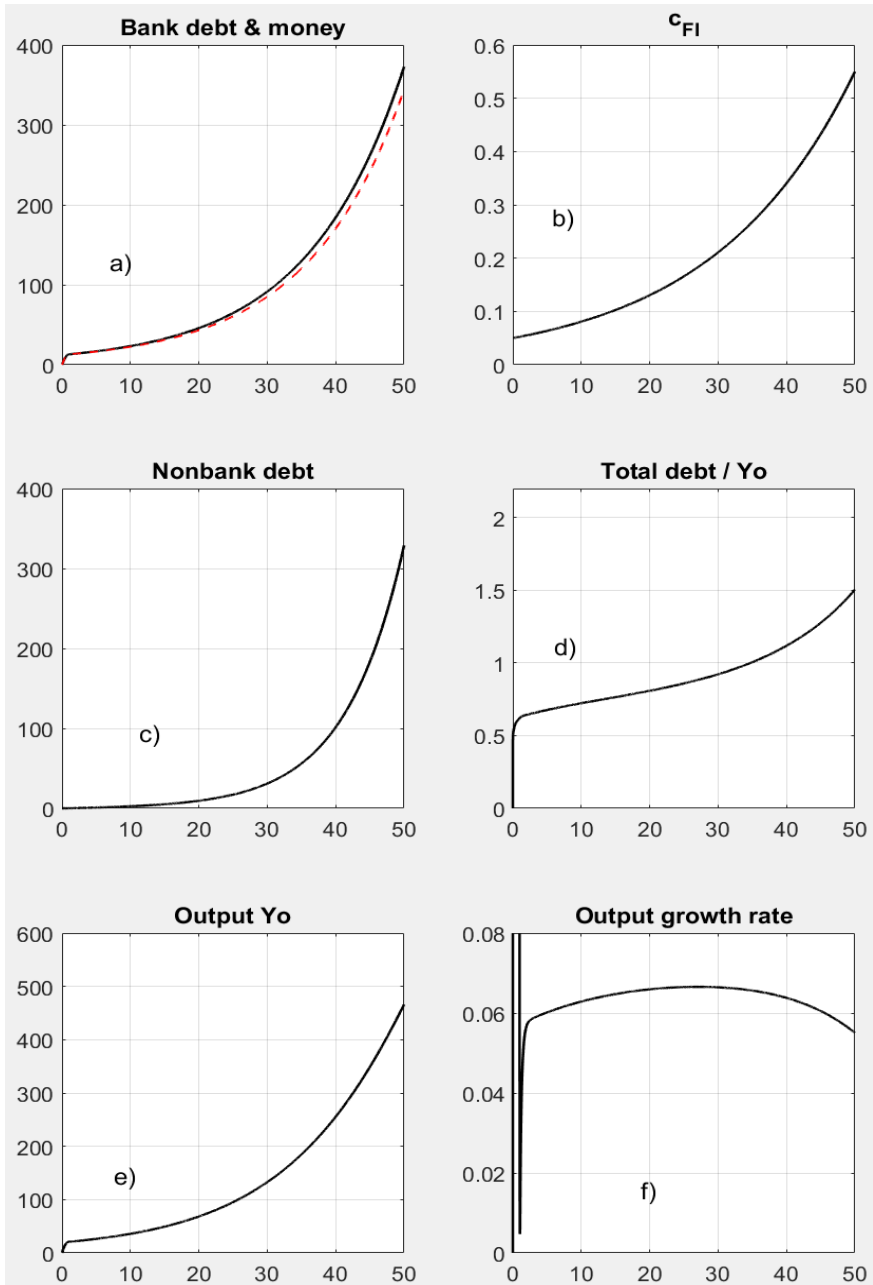


Figure 6.3: Also increasing T_{KF} and π with c_{FI}

Since we assume interest and repayment rates constant¹, we can use the debt burden instead of the earlier introduced debt *service* burden $(i + r)D/Y_o$ introduced in subsection 4.3.7.

Debt service is a *non-discretionary* flow – you can't decide the size of the payment flow, it is decided by the loan contract and you are obliged by it – while consumption flows are discretionary, at least within some fairly flexible bounds (the worker has to eat and pay rent, but he may postpone the purchase of a new TV or holiday). Furthermore, for a capitalist, one may postpone investing in new equipment and buildings, and sack employees or hire fewer of them, and one may reduce own consumption.

When there is persistent relative growth in non-discretionary flows like debt service, the economy as a whole becomes less resilient and more fragile. The first result of this is a rise in the frequency of insolvencies. So now the loss rate λ on loans will no more be constant, but a variable. We have by experimenting ended up with the polynomial relation

$$\lambda(\theta) = 0.01(\theta/1.8)^6 + 0.002 \quad (6.6)$$

which gives $\lambda \approx 0.002$ for the initial debt burden $\theta \approx 0.4$, and $\lambda = 0.01$ for a closer-to-crisis debt burden $\theta = 2$. $\lambda_K = 2.5\lambda$ always. With this choice of parameters, simulation gives a λ that stays close to 0.002 for around 40 years, but increases sharply after that.

Before proceeding, it is useful to consider the simulation results. They are shown in figure 6.4.

There is a collapse before $t = 50$. In graph f) we observe how output growth tapers off. The simulation was stopped at $t = 47$ since the flow of new loans F_{nl} from the Bank became zero and then negative, which is meaningless. At $t = 44$, Y_o culminates, and falls from 252 to 230 (4.8%) in two years. So our crisis may be considered to break out around years 44 to 45.

Both the Bank and non-banks react to the increasing loss rate λ on loans, by reducing their lending. For the Bank this is expressed by an increase in the target capital asset ratio, above the minimum ratio demanded by authorities. Experiments led to

$$\kappa_0(\lambda) = 0.078 + 1.0 \lambda \quad (6.7)$$

¹Obviously, after crisis has broken out, this is not an acceptable simplification. But our purpose is to examine and discuss the run-up to debt crisis and its outbreak, not the ensuing years in deep recession where the policy response is to reduce interest rates strongly.

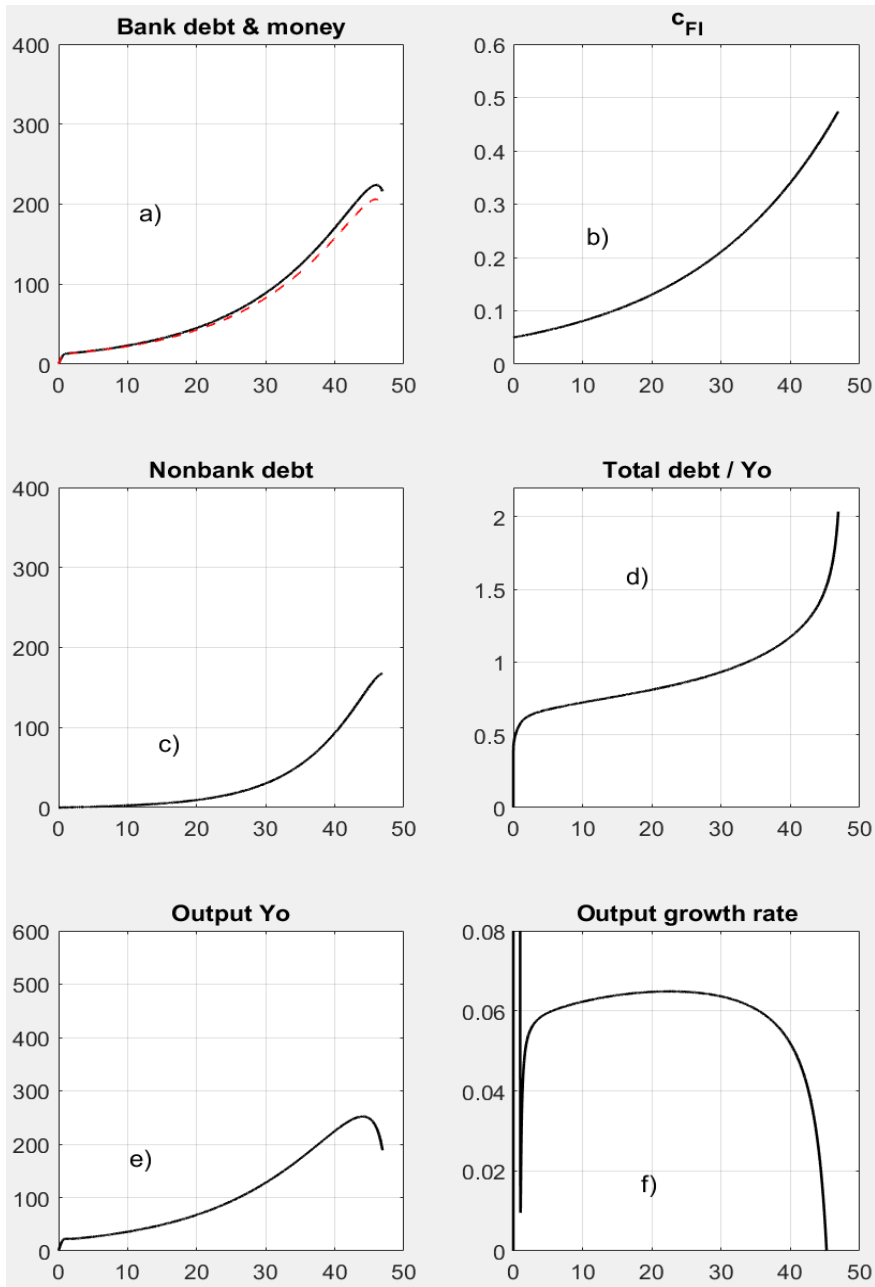


Figure 6.4: Increasing π , all time lags and loss rates \implies collapse

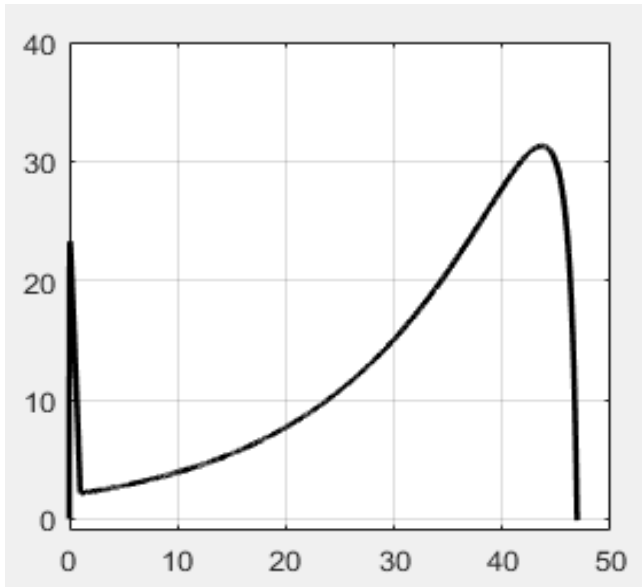


Figure 6.5: Flow F_{nl} of new loans dives to zero

which gives $\kappa_0 = 0.08$ for $\lambda = 0.002$, and $\kappa_0 = 0.082$ for $\lambda = 0.04$, the state at $t = 45$. This seems quite low, but it turned out via many experiments that crisis and collapse became unreasonably dramatic with a steeper increase in $\kappa_0(\lambda)$, so (6.7) was chosen. The PI controller reacts very strongly to the combination of a slight increase in the targeted κ_0 and a small but fast decline in the actual $\kappa(t)$ due to increasing losses. As mentioned, the controller outputs a loan flow which dives to zero, see figure 6.5.

For non-banks, withdrawal from lending means that the average capitalist chooses to hoard cash instead of lending it. This is expressed by an increased T_{KF} as already mentioned. T_{KF} can now increase for *two* different reasons: The first reason has already been introduced in subsection 6.2.2: that the aggregate of non-banks – as financialisation progresses – increasingly lends and transact with others within in the aggregate instead of lending to the real economy. On top of this we now introduce the effect that when crisis is recognised, each non-bank capitalist simply holds back in their lending. Compare the effect of increasing τ in (2.75) or decreasing ρ . Both changes result in an increasing time lag for the aggregate.

We extend the already introduced relation (6.3) from the before-crisis

model, and also introduce the effect of the non-bank loss rate λ_K :

$$T_{KF} = 0.535(0.998 + 100\lambda_K) (1 + 5 c_{FI}(t)) T_{KF0} \quad (6.8)$$

which increases T_{KF} strongly when going from $\lambda_K = 0.005$ to $\lambda_K = 0.05$. At the crisis year $t = 45$, T_{KF} has increased from 0.3[y] to slightly above one year.

Increasing loss rates λ and λ_K give lenders an incentive to hold back. There is however, as already mentioned in subsection 4.3.7, an ameliorating feedback effect of the increasing loss rate. Abstracting from other factors, it decreases debt and thus the debt burden. On the other hand, holding back on further lending reduces demand and thus output, which *cet. par. increases* the debt burden (6.5). So there is a "race" between these factors. But we observe in figure 6.4 that debt reduction because of insolvencies cannot stop crisis and collapse.

We also want to account for the gradual spread of pessimism in the wider population. We do this by introducing a lagged variant of the insolvency (loss) rate. We assume the relation

$$\mu(t) = \frac{1}{1+s} \lambda(t) \quad (6.9)$$

where $\mu(t)$ then is an expression of general pessimism, simply the one-year lagged Bank loan loss rate. Increasing pessimism leads to agents holding back on their non-discretionary flows. For simplicity we assume that capitalists and workers react in the same way. They increase their time lags (decrease their money velocities) following similar equations. Capitalists spending for real investment and consumption is also assumed to be reduced to the same degree when $\mu(t)$ increases. We set

$$T_K(\mu) = 0.55(0.998 + 400\mu)T_{K0} \quad (6.10)$$

Here $T_{K0} = 0.3[y]$, which is the value used for the pre-crisis fixed T_K introduced earlier. Remembering that μ is just a lagged λ , (6.10) increases T_K from 0.3 to 0.82 when going from $\mu = 0.002$ to $\mu = 0.01$, the loss rate around $t = 45$.

For workers we assume

$$T_W(\mu) = 0.83(0.998 + 100 \mu)T_{W0} \quad (6.11)$$

Here $T_{W0} = 0.08[y]$ (= approx. 1 month), which is the value used for the pre-crisis fixed T_W introduced in section 6.1. (6.11) increases T_W from 0.08

to 0.13 when going from $\mu = 0.002$ to $\mu = 0.01$. The less dramatic relative increase in $T_W(\mu)$ is chosen because worker households cannot – because their payments are to a larger degree non-discretionary – hold as much back in their spending as the more well-off capitalists, even if there is a crisis.

To sum up this stage: The dynamics towards crisis follow from growing financialisation as expressed by $c_{FI}(t)$:

- It increases the profit share of output,
- it channels an increasing share of profits to non-bank lending and less to real-economic investment and capitalist consumption,
- this flow increasingly takes a tortuous route inside the non-bank sector before emerging for non-bank loans to the real economy,
- potential lenders (the Bank and non-banks) hold back in their lending when they feel increasing insolvency rates sufficiently,
- and all agents in the economy gradually get pessimistic with a year's time lag or so after the crisis hits the financial sector, and reduce their spending.

This sets in motion a positive feedback process (a "downward spiral").

6.2.4 Holding all time lags except T_{KF} constant

We now wish to let both financialisation and the increase in loss rates and the time lag in the NFF sector increase as in the collapse scenario above, but examine how much better things turn out if firm, worker and capitalist real economy time lags *are not impacted*. The results are shown in figure 6.6.

If we compare this to the collapse scenario shown in figure 6.4, we observe that increasing time lags (decreasing money velocity) as a lagged response in the real economy to an increased insolvency rate, have a strong impact. This confirms what Irving Fisher wrote about the importance of money velocity and pessimism, which we will return to in section 6.3 further below.

On the other hand, comparing the above outcome in figure 6.6 with that in figure 6.3, we observe that this last experiment gives trajectories closer to crisis. That is explained by the loss rate λ being held constant = 0.002 in the experiment that led to figure 6.3.

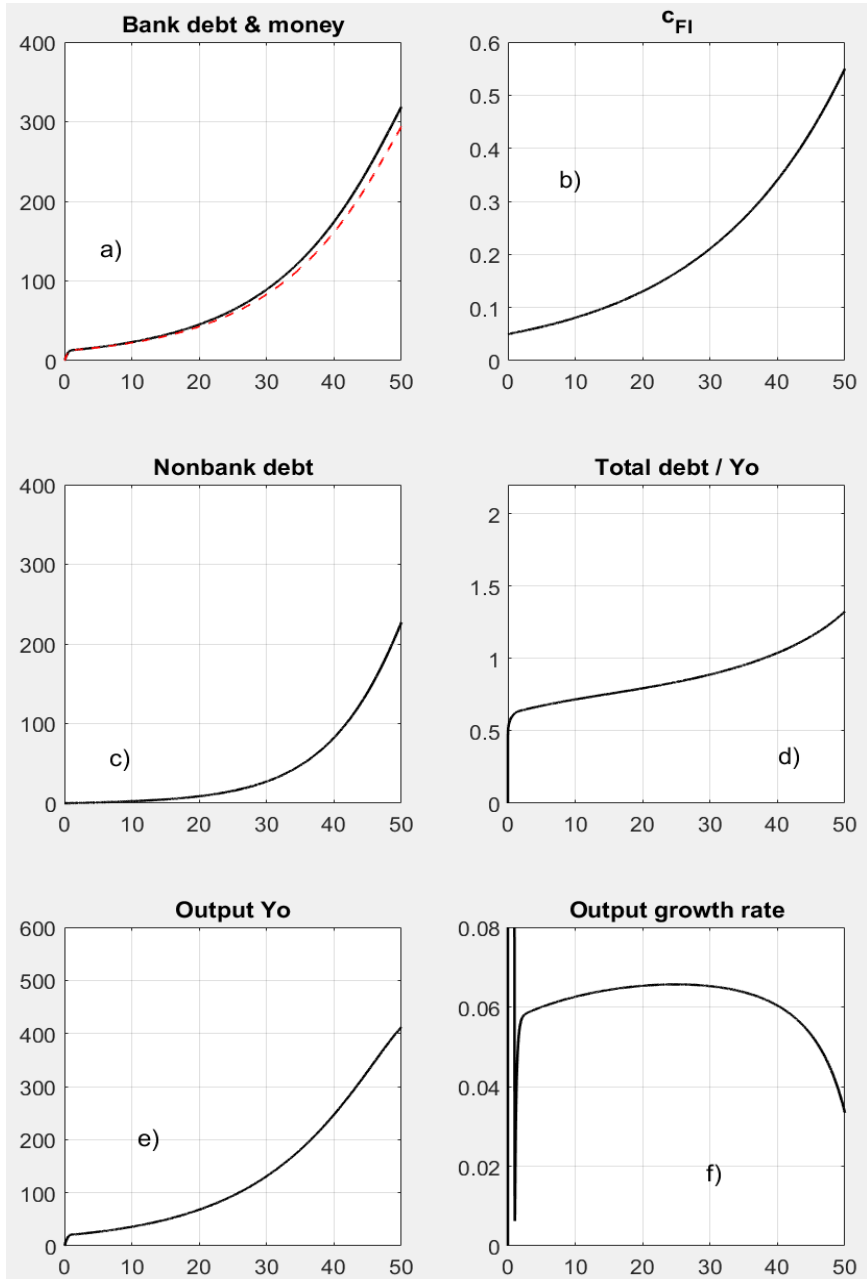


Figure 6.6: Crisis scenario, except constant RE time lags

6.2.5 Constant c_{FI} i.e. unchanging financialisation

We finally wish to examine the effect of (a very hypothetical) *constant* financialisation. We set $c_{FI} = 0.2$, a fairly high value. Capitalists are behaving "financially" from the outset – but *with no change with time in the strength of this behaviour*. All the impacts of other parameters, that led to the crisis displayed in figure 6.4 above, are upheld. The results are shown in figure 6.7. There is now – obviously, following (6.2) – less build-up of non-bank debt. The debt burden converges towards slightly above 100%, and GDP growth rates increases slowly towards 0.07, which is close to the growth rate of money and Bank-issued debt. Output Y_o and Bank debt still grow approximately *pari passu*. Output rises persistently, see graph e). Non-bank debt $D_K(t)$ also (not shown here) converges towards a growth rate of 0.07.

This may seem puzzling, since there now is a debt (D_K) that grows due to "internal" exponential dynamics in a system which additionally is driven by Bank exponential money growth as a system input. But this is a consequence of the external injected money growth rate from the Bank being larger than the internal "inherent" growth rate of D_K given by the non-banks' re-lending (accumulation) loop. More generally stated: a linear system with an internal unstable exponential dynamic with a growth rate b , which is excited by an exponentially growing input a , will exhibit a growth rate converging towards a – and not higher – if $a > b$. (Incidentally, the algebra here is the same as in (6.2), if one interchanges i there with a , and β with b .)

Admittedly our system is only approximately linear, and only so in its first phase along the time axis. We do a last simulation where we increase our constant c_{FI} 50%, to $c_{FI} = 0.3$. The results are shown in figure 6.8.

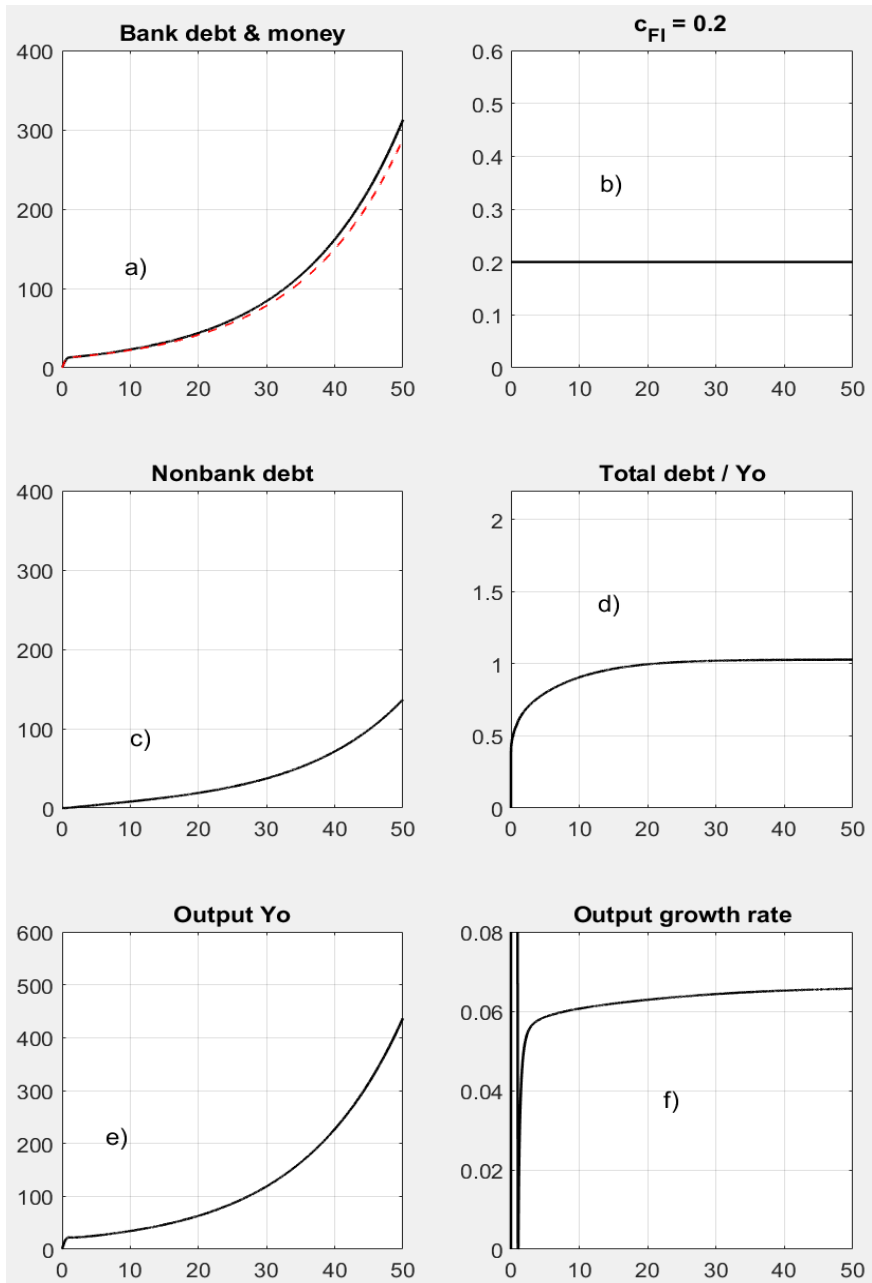
The system collapses, and much earlier than in the crisis shown in figure 6.4 above. So there must be a bifurcation when the constant c_{FI} is changed from 0.2 to 0.3.

Because of much earlier collapse, debt, money and output don't grow to the levels seen in the earlier simulations.

6.2.6 An endogenous c_{FI} instead?

A possible twist on the model would be to not let c_{FI} grow as an exogenous function of time, but decided by feedback, for instance like this:

$$c_{FI} = f \left(\frac{i_K D_k}{i_K D_k + \Pi_R} \right), \quad (6.12)$$

Figure 6.7: Simulation with $C_{FI} = \text{const.} = 0.2$

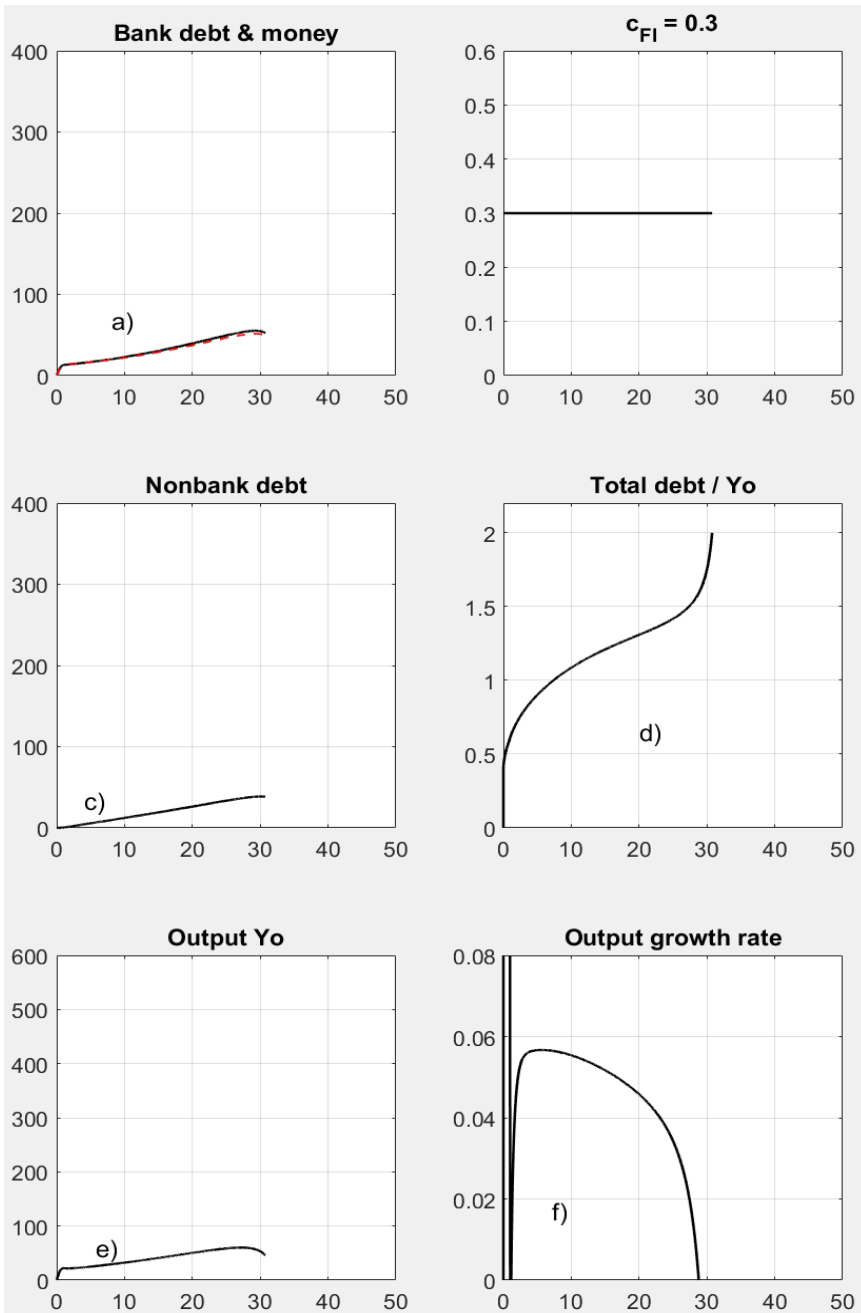


Figure 6.8: Simulation with $c_{FI} = \text{const.} = 0.3$

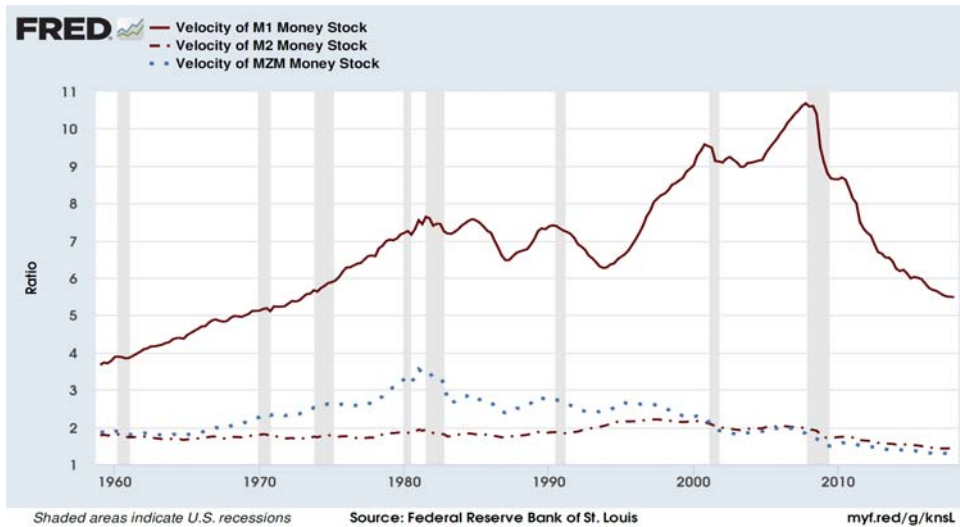


Figure 6.9: Velocity of M1, M2 and MZM in the U.S.

where $f(\cdot)$ is some positive monotonic function. This looks like a reasonable modification. Verbally it says that when capitalists observe that profits from financial investment increase as a share of total profits, perhaps they should allocate (even) more of their income to lending. This would ceteris paribus accelerate the path towards crisis and collapse.

We will not explore this in this thesis however, but conclude the modeling and simulation exercises by stating that the simulation with results shown in figure 6.4 seems to be a reasonable generic representation of what occurs in the real world. It also fits well with Minsky's financial instability hypothesis.

6.2.7 More on money velocity

See figure 6.9: M1 in the U.S. dived sharply and continued falling after the last crisis hit. The fall fits well with the simulations done and the arguments behind our model. But M1 had a long-term velocity growth rate around 2% from 1960 to 2008, in contrast to our model, which assumes constant velocities (expressed by the velocity inverse: time lags of firms, capitalists, workers) over the long-term run-up to the last debt crisis. While M1 is the most important and have grown in that period, MZM has fallen since 1980 and M2 has been largely constant until the crisis arrived around 2008. *But all three measures of velocity have fallen persistently since the last crisis*

started.

We could perhaps – based on the importance of M1 – have assumed a persistent pre-crisis v increase in our model (which uses only a single "generic" measure of money velocity), instead of the simplification with constant v for the pre-crisis period. That said, it is easy to see what a slowly growing v will accomplish. It will *cet. par.* reduce the growth rate of the debt burden as defined in (5.26) since output will increase steeper than commercial (money-creating) bank debt.

To achieve crisis, one may then tweak parameters somewhat to favour steeper non-bank debt growth. This is not explored in this thesis, but is a possible topic for future research.

6.2.8 A simple graphic portrayal of the model

For clarity we will now show a simple graphic portrayal of our model. This is also because we are going to use variations of it in presenting the control solution proposals in the next chapter. See figure 6.10. The symbolism resembles the one shown in figure 4.10. The grey double arrow is just to remind the reader that we have left out the impact of flows between banks and non-banks, mentioned earlier.

We will later build on figure 6.10 to display a circulatory model with an expanded role for the Central Bank and where all money is HPM. This model is shown in the next chapter, figure 7.1.

6.3 Further about debt crisis models

There are a few well-known debt-related crisis models in the literature. One is Irving Fisher's debt deflation mechanism. When crisis has set in, the economy goes into a depressive phase, described by Fisher like this:

Assuming, accordingly, that, at some point of time, a state of over-indebtedness exists, this will tend to lead to liquidation, through the alarm either of debtors or creditors or both. Then we may deduce the following chain of consequences in nine links: (1) Debt liquidation leads to distress selling and to (2) Contraction of deposit currency, as bank loans are paid off, and to a slowing down of velocity of circulation. This contraction of deposits and of their velocity, precipitated by distress selling, causes (3) A fall in the level of prices, in other words, a swelling of the dollar. Assuming, as above stated, that this fall of prices

is not interfered with by reflation or otherwise, there must be (4) A still greater fall in the net worths of business, precipitating bankruptcies and (5) A like fall in profits, which in a "capitalistic," that is, a private-profit society, leads the concerns which are running at a loss to make (6) A reduction in output, in trade and in employment of labor. These losses, bankruptcies, and unemployment, lead to (7) Pessimism and loss of confidence, which in turn lead to (8) Hoarding and slowing down still more the velocity of circulation (Fisher; 1933a).

His emphasis on slow-down of velocity and increasing pessimism is in accordance with central assumptions behind the model presented here. But Fisher's model mostly is about what happens *after* the collapse, while the model here describes the build-up and the collapse, and stopping there. Our model is not valid for the ensuing depressed phase, where Fisher's debt deflation is at work.

Another work is Steve Keen's long-term debt collapse model, where he builds on the Goodwin business cycles model (Goodwin; 1967), (Blatt; 1983). Goodwin's model was a pioneering work, a blow to the erroneous conventional wisdom in economics that cyclical swings in the macroeconomy can only be due to "external shocks". His model exhibits *endogenously* generated cycles. But it contained no financial sector, and no debt.

Keen has extended it to incorporate debt (burden) growth and its interaction with Goodwin-type cyclical dynamics (Keen; 1995). His model exhibits debt-induced collapse for large debt burdens to firms, under reasonable assumptions. But *mechanisms* behind the collapse are different from the model in this thesis. One could envisage a modification and extension of the model developed here, where Goodwin-type "predator-prey" worker/capitalist dynamics that give business cycle oscillations are added. But this is outside the scope of this thesis.

6.3.1 A suggested cycles, bubbles and waves taxonomy

The macroeconomy is cyclical on many time scales, and cycles have been and is a very large field of economic research. A (too) general term often used by economists is "business cycles".

What I miss are attempts at classification and differentiation between bubbles, cycles and other periodic time paths, among other things related to their time scale. I have observed Minsky type analysis being applied to what the speaker/writer often calls "business cycles". In my opinion this is

unsatisfactory.

The following is an attempt to make things clearer:

1. Business cycles have a duration of single-digit years. They may be explained by worker-capitalist struggle over their shares of output (as in the Goodwin cycle model) – and/or time lags in capacity build-up to overcapacity, bust and recession, (Sterman; 2000), ch. 20.
2. But Minskyan analysis relates to much more long-term and slow dynamics (several decades), with a bigger and more dramatic end. The time paths in this analysis do not put weight on the shorter term cycles, but on slowly growing exponential paths that undergo a final collapse. The mechanism may be explained like this: for the financial sector and the public to forget the last severe depressed period (like the 30'ies Great Depression and today's global financial crisis), and for the political atmosphere to change so much that the regulatory system may be weakened, decades are needed (in the order of 40 to 50 years). The same time scale applies to the classical financial accumulation mechanism due to compounding of interest, emphasised by Michael Hudson – see for instance (Hudson; 2010). It needs several decades to make itself dangerously felt, and is closely related to the Minsky process. The compounding mechanism in itself increases the pressure on the economy, while the Minsky mechanism is also psychological/political. I will call this process a *Minsky-Hudson (M-H) wave*. This is a wave that is not easily noted before it is too late, since its growth rate per year is so low, in the order of a few per cent.
3. We also have a third – and in the later "financialised" years more important – category, what should be called a *bubble* and not a cycle, which are booms in asset prices – for instance in property or stock markets, and which differ from category 1. and 2. above. Their time scale is of the same order as business cycles or somewhat longer, but they are not that much related to category 1, worker/capitalist conflict or time lags and overcapacity build-up.

This last category is, however, related to the Minsky-Hudson wave in the sense that when the M-H wave approaches its end, the increased pressure for profitable reinvestment of the huge incoming financial income flows increases recklessness and pressure on authorities for (further) financial deregulation. Deregulation in the next round results in more frequent and dramatic bubble

events. So, imposed on the long M-H wave (type "2") there will be shorter-period bubbles (type "3") with increasing strength as collapse of the long-term M-H wave approaches. And the collapse of the latest bubble may be a triggering factor for the larger crisis that terminates the M-H-wave². This leads to many analysts and commentators focusing on the asset bubble – and by this losing sight of the underlying main cause of the crisis, the slowly but mercilessly growing M-H indebtedness wave.

Summing up, a (hopefully) useful taxonomy and terms are proposed, listed by their degree of severity:

- *Business cycle,*
- *Asset bubble,*
- *Minsky-Hudson wave.*

²I generalise here, perhaps to much, because one might argue that the collapse of Soviet Union around 1990 weakened the socialist left and the unions. This changed the power balance and partly explains the increased frequency of speculative bubbles at the end of the current M-H wave.

Chapter 7

Reform proposals

As initially stated, this author and thesis has one foot in the control systems community and one foot in macroeconomics. Over the years I have observed a certain difference between the social sciences and engineering sub-cultures. Academics in the social sciences (economists included) seem often to be satisfied with *describing* problems, while engineers describe problems or challenges with the intention of *solving* them.

In the spirit of engineers – and to be fair I have to include Karl Marx¹, Irving Fisher in *Stamp Scrip ch. VI*, and the MMT economists – here follows a chapter with solution proposals. This chapter mainly builds on three papers, (Andresen; 2013),(Andresen; 2014) and (Andresen; 2012).

7.1 The two unsolved problems

The main unsolved problems in avoiding debt crisis, examined in the previous chapter are:

1. Persistent debt/GDP growth leading to crisis, and ...
2. crisis and collapse finally occurring, leading to years of depression / deep recession.

When discussing how to respond, one should distinguish between these two.

¹"Philosophers have hitherto only interpreted the world in various ways; the point is to change it"

- Karl Marx, Eleven Theses on Feuerbach

7.1.1 Avoiding debt burden growth

The best solution is of course to avoid the first problem altogether: Growth in D/Y_o to dangerous heights. The primary measure here is higher taxation rates on non-banks and curbing related speculative activity there. In the model in figure 6.1, increased tax rates correspond to a lower equivalent (after taxes) interest (or return) rate i_K , leading to the intended weaker D/Y_o growth. One should probably aim for a constant D/Y_o ratio – there is no rationale for persistent growth in this ratio.

Commercial (licensed) banks need – theoretically – not be included in such a stricter tax regime, since they create money at the same growth rate as debt. One could make non-bank lending so unattractive (also through different regulation measures, not only higher tax rates) that it was largely eradicated. The alternative for would-be investors would then be investing in real-economy firms or saving in commercial banks. Increased savings with commercial banks would not change their debt and money creation rates, since their solvency (capital-asset ratios) would largely be unchanged (following the analysis in the previous two chapters).

The remaining issue and potential problem would be the possible inflationary effect of money creation at the same rate as debt. Then real growth has to keep up (and we have the issue of environmental sustainability, which – however – is not a topic for this thesis). But inflation is a possible problem also in today's regulatory environment with endogenous credit money growth.

That said, inflation may actually be controlled much more effectively, and we will present a new idea on this later in subsection 7.3.3.

7.1.2 How to get out of a debt-induced crisis?

One also needs to have recipes for what to do if a debt-induced crisis has arrived because D/Y_o has been allowed to grow to a dangerous level. A method that has been used much during the global debt-induced deep recessions since 2008 is "Quantitative Easing" ("QE"), where Central Banks buy "rotten" financial paper from banks to avoid insolvencies there. This puts banks on a much better footing concerning their capital/asset ratios – (hopefully) giving them an incentive to invest in the real economy to get the economy moving and reducing unemployment. But even if banks become solvent through QE, this cannot *force* them to invest in the real economy. This is like pushing on a string. Even if they have received HPM for rotten paper, they may still hold back in lending to the real economy. It turns

out they have, to a problematic degree, instead chosen to speculate in asset markets.

Another tool that has been extensively used is reduced interest rates, which are basically controlled by the Central Banks. But interest rates cannot go below zero, where they have been held near for many years now (per 2018). And interest rates close to zero is *cet. par.* an incentive for banks to speculate in asset markets because real-economy loans are not very profitable.

We will use the rest of this chapter to describe and argue for tools and reform ideas that are more potent than today's, also for getting out of a depression-like crisis. *These tools and reforms presuppose the advent of (mostly) cashless, electronic money.* Luckily, what seems to be unstoppable developments in this direction are occurring quite fast worldwide – also in poorer countries.

7.2 Possibilities with an electronic money system

Physical currency is being phased out as an important means of exchange both in developed and developing countries. Today it is technically feasible to discard bills and coins completely and do all transactions by debit card, personal computers (both quite common in developed countries), and/or via the mobile phone network. Mobile phone money transfers have a proven track record especially in poor countries, for instance the pioneering "M-Pesa" in Kenya (Hughes and Lonie; 2007). With electronic money ("EM") all transactions are reflected in movements between accounts in the involved banks.

In the proposed implementation here however, there are no deposits with private banks: All accounts are at the Central Bank.

All citizens and firms are offered EM accounts at the CB. The advantages are obvious and many:

1. The system is very cheap to run, compared to a system with bills and coins.
2. Adjustments that turn out to be needed can be implemented in software, therefore very easily and cheaply. No cumbersome and expensive printing/stamping and distribution of bills and coins.
3. Forgery is impossible. So are robberies.

4. This is a 100% reserve system. All deposits are HPM (base money), at the CB. No deposit insurance needed. Money cannot be lost as argued earlier. This is clear to the public – *no bank runs*.
5. EM is an extremely inclusive and convenient system, giving rural sectors of a poorer country – where ATMs and bank branches may be far between and not all people have accounts – a tool for easy economic participation and exchange.
6. A black economy in EM close to impossible. The same holds for tax evasion. Intelligent software can monitor transactions 24/7, and flag human operators when suspicious patterns emerge. Knowledge of this implies a credible threat, so that agents to a significant degree will abstain.
7. EM cannot be used for capital flight, since it only resides at the CB. All foreign transactions are logged and thus controllable, as mentioned in the previous point.

Finally, two unconventional advantages/possibilities:

8. Negative interest on money held ("demurrage") may be easily implemented, to speed up circulation if that is needed.
9. A new possible control tool with the opposite effect is feasible by money only existing as accounts at the CB: A tiny but adjustable transfer tax between any accounts.

We will return to these two tools further below.

EM is taking over for bills and coins. This is technologically driven. It is not something one can decide to abstain from, it is unstoppable. The technological development process that allows electronic transactions instead of exchanges using physical currency, has the same merciless and irreversible character as the advent of the electronic calculator in the 70s and digital photography in the 90s: it meant the unavoidable death of the slide rule (then) and photographic film (more recently). Based on the nature of technological innovations and the market economy's exploitation of such, we may predict the death of physical currency. It is a question of when, not if, this will take place.

So we should discuss how to adapt – and of course how to exploit the possibilities that such systems give for better monetary and economic control. These developments enable some very useful implementations, among

these new and – for society – beneficial roles for the Central Bank. We assume a scenario where the country in question issues its own currency, and all money is “electronic” – no bills and coins. We also assume that all deposits are at the Central Bank – designated “sovereign money”.

7.2.1 The Central Bank and electronic money

The role of a CB has up to this day been as an interest-rate setter behind the scenes and – in crisis – “lender of last resort” for the network of private licensed (“commercial”) banks. The public has had no relationship with the Central Bank, but only with commercial banks. The commercial bank network has historically been quite dense, with branches of competing banks within a reasonable distance from customers. The reasons for this geographical diversity has been twofold:

1. Handling deposit accounts and receiving or furnishing customers with physical currency.
2. Vetting potential borrowers and extending loans.

With the advent of electronic transactions (via PC, debit card and mobile phone) the need for a dense network of branches has decreased, and commercial banks have started the process of closing down an increasing share of these. If we envisage an expected future without physical currency, the first point above will disappear as a reason for having bank branches. What remains is the second point, the need for offices to handle decisions about loan applications, which to a fair degree will be best handled by personnel having local and/or specialised knowledge. Except for this, most decisions can be made at a bank’s central office.

So, simply because of no more need for branches to acquire or deposit physical money – for purely technological reasons, not society’s economic policy considerations – it becomes feasible for all “agents” (persons, firms) to only have their checking accounts directly at the CB. Then one may dispense with bank credit money and let all money in circulation be base money (high-powered money; HPM). For the public this means that their deposits are completely safe, and in that sense it matches the 1930ies “Chicago plan” and its “100% money” proposal. But it goes further, because in the 100% money plan, banks would hold people’s deposits (although fully covered by the banks’ deposits at the CB) and furnish them with physical currency, while in the above electronic money scenario liquid deposits only exist directly at the CB. This is the *sovereign money* alternative (Huber; 2017). Technological

possibilities today give a new impetus to this alternative, which has until now only been promoted based on political economy arguments.

We will from now on describe and argue for a banking scenario based on the sovereign money alternative, but with only electronic currency, using the acronym “ESMA – Electronic Sovereign Money Alternative”.

We will first address an argument raised against both 100% money and sovereign money: “*when banks are not allowed to create credit money, the economy and employment will suffer because of lack of credit for investment*”.

7.2.2 The "heterodoxy" is partly hostile to 100% money

The famous pre-WWII Fisher et al Chicago Plan, was more recently re-examined (Benes and Kumhof; 2012), and in conclusion supported. They write in the abstract:

At the height of the Great Depression a number of leading U.S. economists advanced a proposal for monetary reform that became known as the Chicago Plan. It envisaged the separation of the monetary and credit functions of the banking system, by requiring 100% reserve backing for deposits. Irving Fisher [reference in paper] claimed the following advantages for this plan: (1) Much better control of a major source of business cycle fluctuations, sudden increases and contractions of bank credit and of the supply of bank-created money. (2) Complete elimination of bank runs. (3) Dramatic reduction of the (net) public debt. (4) Dramatic reduction of private debt, as money creation no longer requires simultaneous debt creation. We study these claims by embedding a comprehensive and carefully calibrated model of the banking system in a DSGE model of the U.S. economy. We find support for all four of Fisher’s claims. Furthermore, output gains approach 10 percent, and steady state inflation can drop to zero without posing problems for the conduct of monetary policy.

Ann Pettifor disagrees (Pettifor; 2013), and argues (p. 20) that 100% reserve banking will lead to lack of credit:

The Kumhof and Benes proposal is indeed based on the monetarist ideas of the Chicago School, one that seeks to limit the quantity of money, and that would restore the role of banks to

intermediaries between savers and borrowers. Only now the proposal is to eclipse the role of the private sector altogether, and only allow lending backed by a 100% reserve requirement. In other words, all banks or lenders would first have to mobilise 100% of the funds needed for lending. This would massively constrain the availability of credit. (...)

Limiting the quantity of credit is certainly one way of limiting employment. Thus monetarist theory and policies both tolerated and sustained a massive rise in unemployment in the 1930s and 1980s. The Kumhof and Benes proposal is no more than a revival of these policies: the ‘barbaric relic’ that was the gold standard.

Pettifor is hostile to the 100% reserve concept – and her platform is anti-neoliberal. She is not alone in this; many central authors in the heterodox Modern Mone(tar)y Theory (MMT) and/or Post Keynesian camps share her position. One of these is Jan Kregel, who describes and supports Hyman Minsky’s critique of what he termed “narrow banking” (this corresponds to banks subjected to a 100% reserve requirement) in a paper that argues along similar lines (Kregel; 2012):

In the absence of a large government sector to support incomes, liabilities used to finance investment could not be validated in a narrow bank holding company structure. But, even more important, it would be impossible in such a system for banks to act as the handmaiden to innovation and creative destruction by providing entrepreneurs the purchasing power necessary for them to appropriate the assets required for their innovative investments.

Emphasising the need for easy access to credit for “Schumpeterian creative destruction”, Kregel argues that if banks are not allowed to create extra money when lending, what is left:

... is not a bank, but simply a safe house or piggy bank for government issues of coin and currency.

Kregel, however, points to a possible solution to – or amelioration of – lack of capital for investment:

In the absence of private sector “liquidity” creation, the Central Bank would have to provide financing for private sector investment trust liabilities, or a government development bank

could finance innovation through the issue of debt monetized by the Central Bank. (...) such a system would have to combine Keynes's idea of the "socialisation of investment" with the "socialisation" of the transactions-and-payments system. (...)

the real problem that must be solved lies in the way that regulation governs the provision of liquidity in the financial system.

We will now describe regulation that may achieve what Kregel suggests, but by the unconventional road of *letting commercial (licensed, "narrow") banks decide on creation of HPM at the Central Bank*. This will be explained in the following.

Ample credit lines for banks at the Central Bank

Consider an economy where all money is base money (HPM): Could banks not – if they trust they have a worthwhile and fairly safe lending opportunity – just borrow HPM from the Central Bank and re-lend it at a somewhat higher interest rate? This is in contrast to today's state of affairs where credit money is created directly through bank lending, completely dominating money growth.

Such bank borrowing from the CB implies that HPM will grow as an effect of this, not only through government deficit spending (if we follow the MMT advice of financing government deficits by directly "borrowing" from its CB, instead of selling bonds to banks and the public). In such a scenario, some money will not only be spent into circulation (government deficit), some will also be lent into circulation (via banks). *But all of it will be HPM*: The amount of extra money created and subsequently put into circulation due to bank borrowing from the CB will be completely safe, not credit money which carries some risk for the owner.

Banks could also gather money for their lending by selling bonds or offer time deposits to the public. Then the amount of extra money created by bank borrowing from the CB would constitute only a share of new loans given. But we argue that this alternative, which puts the saver at some risk, is not necessary.

To sum up at this stage, our ESMA scenario assumes that all government deficits are financed by "loans" from the CB ("loans" in quotation marks, if we regard the CB a tool of the government – following MMT), and that all bank lending is financed by corresponding loans to banks from the CB. Banks will resemble "franchisees" of the CB, living off the difference on interest rates in and out. The role left for banks is to be pure intermediaries.

7.2.3 Ensuring that only bank owners take the hit

With ESMA, if a bank defaults, the bank owners would lose all their equity. *Only the bank owners, not tax payers or society as such, take the hit.* Hence, any bank would have a much stronger incentive for responsible behaviour than in today's environment. And there would be no need or reason whatsoever for society to step in with bailouts. This will be a credible threat that banks' owners have to take extremely seriously.

The impact on the CB would be much less dramatic. The only "loss" to the CB and society when a bank defaulted on loans from the CB, would be that the corresponding HPM that was supposed to be destroyed though repayments with interest to the CB, remained in circulation. Instead of society increasing taxes to pay for bailing out a bank, money that was already in circulation would not be retired. The effect of this would then be spread thinly over society as a whole. At worst this would give a small impulse for inflation. We may contrast this small disadvantage with the big advantage of directly hitting the bank owners in this proposal; they would lose their assets. And this no-need-for-bailouts system should – even with interbank lending – be 100% robust in a systemic sense.

Based on the above, licensed banks can be given very ample credit lines to the CB, at reasonably low rates. This is the main reason that society's need for credit would not be constrained in a damaging way, as Pettifor fears.

That said, banks should not be allowed unlimited borrowing from the CB. This is discussed in the next subsection.

7.2.4 A BIS-type capital adequacy constraint still feasible

In today's regulatory environment, banks are to a decreasing degree reserve constrained in different countries. The trend is towards implementing Bank for International Settlements-type regulation that only sets a lower threshold for commercial banks' capital adequacy. We will discuss this based on the simplified representation of this regulatory framework that has been used in earlier chapters, and in the next stage suggest a very similar capital adequacy requirement tailored for the ESMA scenario.

We define:

$M(t)$ = deposits = money stock [\$], a bank's liability.

$D(t)$ = loans from the bank [\$].

$R(t)$ = reserves = a bank's deposit with the CB [\$].

The Bank's total financial assets are $D + R$.

κ_0 = the required minimum capital/asset ratio []. This entity is dimensionless, thus the [].

κ = the actual capital/asset ratio [].

Variables' dependence on time (t) is from now on not shown. Remembering the Basel rule that risk weights shall only apply in the denominator and that reserves R carry zero risk weight, we have

$$\kappa = \frac{D + R - M}{D}, \text{ and get the requirement } \kappa \geq \kappa_0 \quad (7.1)$$

What happens when the bank extends a loan ΔD ?

Since both D (the bank's asset side) and M (liabilities side) increase with ΔD , the numerator remains the same, while the denominator increases with ΔD . The result is a fall in κ towards κ_0 . This may imply a restriction on further lending, and that is the purpose of the regulation: to achieve some minimum robustness against insolvency.

Now to the ESMA scenario. We additionally define:

D_{CB} = the bank's debt to the CB [].

D = loans from the bank as above [].

We now have for the bank,

$$\kappa = \frac{D + R - D_{CB}}{D}, \text{ demanding } \kappa \geq \kappa_0 \quad (7.2)$$

The liability towards the CB here plays the same role as today's bank liability M towards its depositors. When the bank extends a loan ΔD , the situation for the bank afterwards is:

$$\kappa = \frac{D + \Delta D + R - (D_{CB} + \Delta D)}{D + \Delta D} \quad (7.3)$$

The ratio κ is decreased in the same way as today. We thus achieve a regulatory constraint on banks' lending behaviour which formally is quite similar to that given by the current regulations, except for one difference: in today's environment we have a two-way relationship (the bank and the borrower), while the relationship in the ESMA scenario becomes triangular: the bank, the CB, and the borrower. The bank gets a claim ΔD on its borrowing customer like today, but now the CB gets the same extra claim on the bank. The balance sheet of the customer at the bank increases with ΔD on his liability side, while the customer's checking account at the CB is credited with ΔD .

The change from a today's dual to this triangular relationship is shown in 7.1, lower right. Compare it to today's system, earlier shown in figure 6.10

and for convenience repeated here, upper left. Now all money is HPM and cannot be lost. Capitalists (or their behaviour) are divided into two types. Type "II" invests financially with the CB, which offers a choice of paper, interest rates and maturities – also to banks and the wider public. Note that direct investments in firms are not part of the flow structure; these flows occur *inside* the real economy.

Note also the double arrow from the CB into the real economy. One signifies injection of HPM via Bank lending. The other is the government's deficit spending – not accompanied by a corresponding debt increase. Money is spent, not lent, into the economy. This second flow means that money growth can be steeper than debt growth, in contrast to today's state of affairs, analysed in chapter 5. And all money is HPM, not credit money.

7.2.5 The CB as a lending and savings hub for society

So far about CB lending to banks. On the savings side, the CB can – due to the information technology revolution discussed above – offer individual accounts not only for banks, but for all agents: citizens and firms; both a checking account and a spectrum of time deposits yielding different rates, payment profiles and durations. Since individual depositors' money at the CB – whether from persons or businesses – would be completely risk-free, a checking account there should yield zero interest. Such accounts could be cost-free for the user, considered part of a modern welfare state's shared free infrastructure, like healthcare and schools.

By tweaking interest rates on its lending, the CB can ensure that banks get the necessary incentives to borrow and lend, by a sufficient difference between bank lending rates and (for them) CB borrowing rates. On the other side, by offering sufficient rates on its spectrum of time deposits and other instruments, the CB can withdraw money from aggregate demand – from banks, firms and individuals. This largely removes the need for the bond markets we have today, and by this the damaging power they hold over the political-economic choices that are made.

7.3 Inflation and deflation control

7.3.1 Electronic money and MMT

In this author's opinion, the best theoretical platform for the understanding of today's macroeconomies and what might be done to improve them, is Modern Mone(tar)y Theory – MMT – also labeled "neo-chartalism". MMT

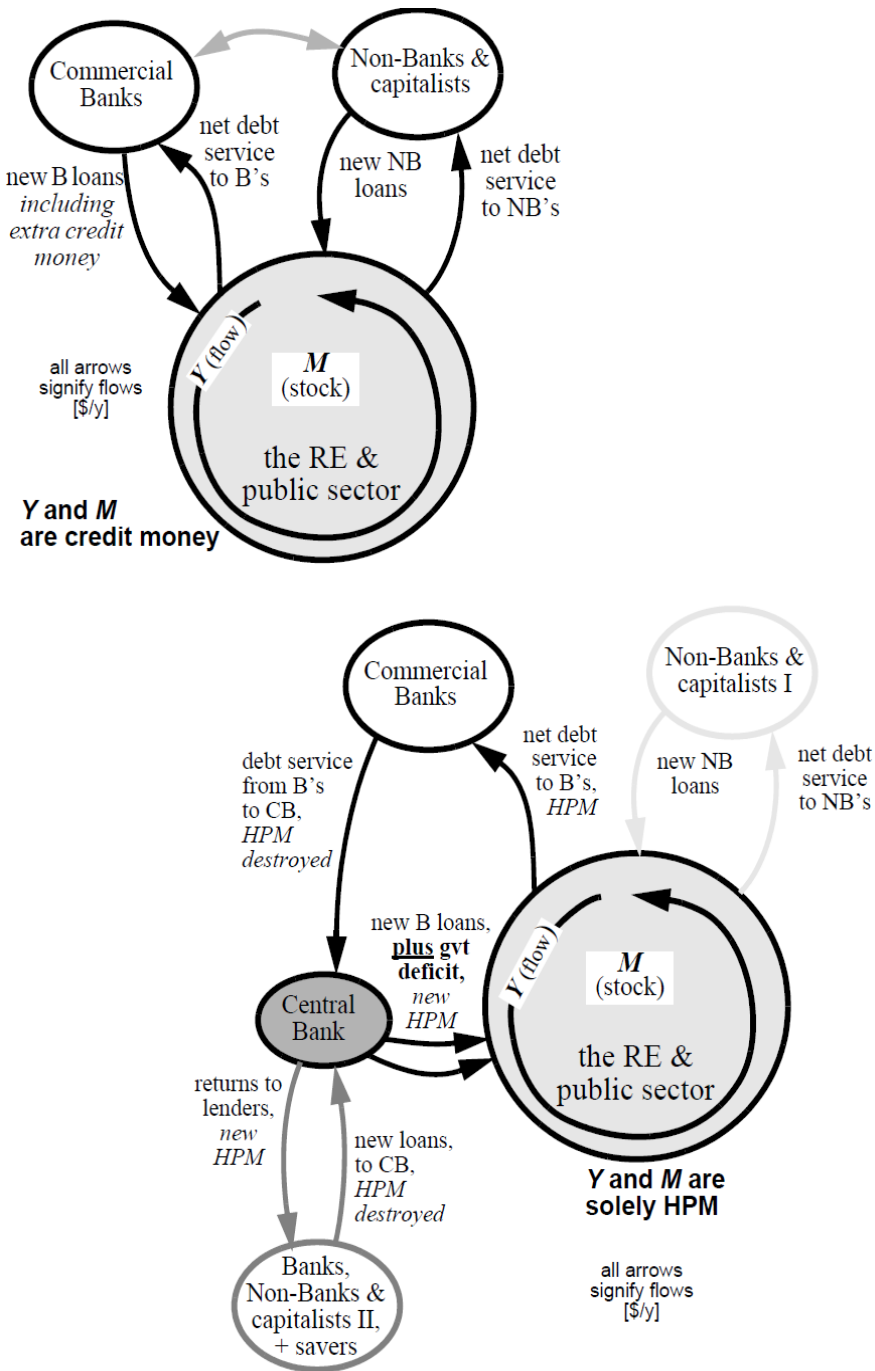


Figure 7.1: Flow diagrams of today's and a HPM-based economy

has since the onset of the debt crises around 2008 gained influence in the global discourse on macroeconomic theory and crisis solutions. A comprehensive text explaining MMT is (Wray; 2012).

In the MMT framework, the government and the Central Bank is seen as one unit. The "independence" of CB's that is the rule in most countries is a political and legal construct, and may as such be reversed by a national assembly. Any CB is constitutionally, at least in some final instance, an arm of the government. This is generally accepted, not solely by MMT adherents. For a country issuing its own currency (this is a prerequisite for MMT to be valid as a platform for policy), a government's "debt" that builds up with its CB through deficit spending in excess of the income from selling bonds, is only an accounting convention. A government does not need to "finance" its spending through tax income or to borrow by issuing government bonds – a government may spend (and thus net create money) by debiting its account at the CB. Such a government is not revenue constrained. It can never "run out of" its own issued currency, and can always pay any debt if this debt is in its own – not foreign – currency. The role of taxes in MMT is to give money value, drain money to control demand and limit possible inflation, and to redistribute income. In the MMT view, money has value and enjoys confidence since it is the only accepted means to pay taxes, and since the state can enforce tax payment. It does not need to be backed by any asset.

MMT assumes flexible exchange rates. Rigidly binding one's currency to foreign currency(-ies) removes the advantages of MMT: one is then on a de facto "gold standard", and this is incompatible with MMT.

So far on the main characteristics of MMT.

The obvious and common objection to MMT is "it will be inflationary". Yes, inflation may be an issue. This is a reasonable objection and will therefore be discussed below. That said, inflation is a possibility under any macroeconomic regime if aggregate demand is near or surpasses some capacity limit. *The possibility of inflation is not in itself an argument against MMT.* Through taxation and other methods, inflationary pressures can effectively be taken care of within an MMT paradigm. How to achieve this is one of the topics below.

As discussed, a government may use the option of injecting new fiat money (base money – HPM) into circulation. But in today's system we have net creation of money through bank lending. This credit money – as already argued – grows endogenously. Control of money supply from the CB, as told in the monetarist and mainstream economics money multiplier story, is not possible.

That said, control of money supply is not the central point in this paper – it will focus on control of another entity: *money velocity*. Control of velocity is much more effective, and it becomes feasible – for the first time in the history of money – with electronic money and no physical currency.

In a recession or even depression-like situation – the case in many countries today – the attraction of MMT is obvious: since a government with own-issued currency is not financially constrained, such a crisis can be remedied by running arbitrarily high fiscal deficits as long as needed, i.e. spending extra HPM into the economy to employ people and buy goods and services. A government issuing its own currency can always employ all the unemployed.

But there is a challenge to MMT that has not been much discussed by its proponents: in the opposite scenario, if an economy is running close to full capacity or beyond (for instance after a crisis where a large amount of money was injected, remaining in circulation), and there are ensuing inflationary pressures: how can a government restrict money flows? This is a genuine problem, and is not easily solved in today's technical monetary environment. But there are solutions to this if all circulating money is electronic.

7.3.2 A problem – injection and drainage asymmetry

There will be negligible opposition in a depressed situation if a government hires more people and buys more goods and services, with brand new HPM, created *ex nihilo* at the CB. Such policy becomes possible with an MMT understanding of macroeconomics. And in an economically depressed situation, people will gratefully accept this, in spite of alarms from deficit hawks and many financial pages pundits.

But when a government tries to drain money back later on in a boom, running a surplus over time through increased taxes, there will probably be strong popular resistance. Furthermore, in a boom there will usually also be a widespread over-optimistic mood in the population, enhancing such resistance – which can take many forms: media campaigns, demonstrations, capital flight, tax avoidance, stashing away cash, voting for parties arguing for "small government" with low taxation.

MMT proponents have to address this issue, even if this is a hypothetical scenario diametrically opposite to today's (2018). For it is difficult to convince the public, academics and decision makers today of the acceptability of large and persistent (over years) deficit spending, if one does not have a recipe for what to do in a later boom. Paul Krugman writes:

It's true that printing money isn't at all inflationary under current conditions— that is, with the economy depressed and interest rates up against the zero lower bound. But eventually these conditions will end. At that point, to prevent a sharp rise in inflation the Fed will want to pull back much of the monetary base it created in response to the crisis, which means selling off the Federal debt it bought. So even though right now that debt is just a claim by one more or less governmental agency on another governmental agency, it will eventually turn into debt held by the public (Krugman; 2013).

7.3.3 M and v control

It is first necessary to make an important point about money supply and money flows. Demand in an economy is not decided by the aggregate money supply (a *stock* [\$]), but by the aggregate Y of money *flows* [\$/y]. Using nominal entities we have $Y(t) = M(t)v(t)$, where M is aggregate money stock and v is average money velocity. This is the quantity equation, espoused by monetarists, and (much for the same reason) derided by the economics heterodoxy. In this author's opinion, the monetarists are wrong because they ignore v and focus solely on M . There are also mainstream economists who point to the insufficiency of using M as a control variable:

In terms of the quantity theory of money, we may say that the velocity of circulation of money does not remain constant. "You can lead a horse to water, but you can't make him drink." You can force money on the system in exchange for government bonds, its close money substitute; but you can't make the money circulate against new goods and new jobs (Samuelson; 1948).

But many in the heterodoxy are also wrong – not because they (correctly) argue that M is not a sufficient control variable – but because they consider v of no importance:

Unfortunately, most economists are brainwashed with the trivializing formula $MV = PT$. The idea is that more money (M) increases "prices" (P) – presumably consumer prices and wages. (One can ignore velocity, " V ," which is merely a tautological residual.) " T " is "transactions," for GDP, sometimes called " O " for Output (Hudson; 2010).

This might be characterised as throwing the Mv baby out with the M bathwater.

In contrast to this, Irving Fisher really recognised the importance of velocity:

Free money may turn out to be the best regulator of the velocity of circulation of money, which is the most confusing element in the stabilization of the price level. Applied correctly it could in fact haul us out of the crisis in a few weeks ... I am a humble servant of the merchant Gesell (Fisher; 1933b).

Fisher argued for a parallel money in the depression-ridden U.S. (Fisher; 1933b), and levying a holding fee (negative interest, demurrage – originally proposed by the German-Argentinian merchant and monetary theorist Silvio Gesell) on this money to force agents to spend. *Thus it would be possible to increase activity even for a small M , due to higher v .*

Fisher understood that v is not a "residual" as Hudson calls it, but an important behavioural variable, and that it would be low in a depression, and needed to be boosted. It is strange that this is not more recognised, since v is in a one-to-one relation to (inverse) liquidity preference, and liquidity preference is a concept that is widely accepted and used among macroeconomists – not the least by Post Keynesians, who are very much against the quantity theory.

With electronic money one is able to not only enhance control of M , but also achieve control of v , which until now has been mostly ignored (among several reasons because such control is very difficult in a system containing physical currency). While M cannot be changed significantly within a short time span (since it is a stock and needs time to change, and since draining M will be a politically controversial extra tax), this may be done with v (since it is a behavioural variable not a stock, and no liquid assets are taken from the holders). By having control of both M and (especially) v , one may exercise potent control of their product, $Y = Mv$.

There are (theoretically) a quadruple of ways to do $Y = Mv$ control:

1. A fee (negative interest, *demurrage*) on money held: M decreases slowly, v increases strongly and immediately, therefore Y increases immediately. And a government can exploit shrinking M by creating a corresponding extra HPM flow and thus spend more. This is a bonus in a recession/depression.

2. A fee on transferring money between accounts: M falls slowly, v falls stronger and immediately, therefore Y decreases immediately.
3. Positive interest on deposit money, the opposite of item 1. This is today's sole tool: M increases slowly, v may decrease a little but slowly, therefore Y hopefully decreasing, but this is very mood-dependent.
4. A small reward for transferring money between accounts: the opposite of item 2, v increases strongly, M grows persistently, Y "explodes".

Item 4 is obviously absurd, since agents can then increase their money holdings just by transferring money back and forth. It will be ignored in the following.

I will now discuss the new possibilities given by items 1 and 2, and especially item 2. Negative interest on money held (item 1) works, as demonstrated by local crisis parallel currency issued in the Austrian town Wörgl in 1932 (Lietaer; 2010), where money velocity turned out to be 12 – 14 times the velocity of the Austrian schilling . This was also an inspiration for Irving Fisher's (futile) attempts to get a similar solution implemented in the depression-ridden U.S. But the Wörgl technical demurrage solution was cumbersome: one had to buy a stamp every month and glue it to a bill, in order for the bill to uphold its validity. And with coins one cannot even do that. With electronic money however, it is exceedingly simple: every day a tiny proportion of the amount in a checking account is deducted. And this proportion may be easily adjusted as the state of the economy changes.

Now to item 2: a fee on transferring money between accounts. As far as this author knows, this is a new concept – easily implemented in an electronic money framework – that has not been considered in the large economics literature on inflation control. One could object that it resembles a value added tax, but the important difference is that the fee is on *all* transfers, not only for purchases from firms (one should of course have a VAT like today, in parallel with an account transfer fee). This property, combined with all money residing as checking accounts at the CB, makes avoiding the fee impossible and removes all need of human control. The size needed for such a fee to have an impact is difficult to decide *ex ante*, but a conjecture is that this measure will be quite potent. One could start with a very low (and therefore economically and politically harmless) level – say 0.1% – and monitor the impact. If the impact in a trial period is too small, increase the fee a little.

This type of fee would be incredibly more effective to damp an overheated economy, than today's blunt tool of a CB interest rate increase. It can stop

too much spending in its tracks.

7.3.4 Fiscal policy and capital controls

From an MMT perspective, fiscal policy is more important than monetary policy. All money as electronic HPM in accounts at the CB will make taxation and levying of fees easier. This will be the case both for collection, control and adjustment. Tax evasion and crime will be sharply reduced as already mentioned. The need for human control will be much lower, since detailed monitoring may be done by software which alerts human operators only when suspicious patterns are detected.

Possibilities for capital flight will be sharply reduced, even if this cannot be completely eradicated (capital controls in an electronic money environment should be a topic for further research).

Electronic money, applied with an MMT understanding, enables a revolution in macroeconomic control. But this insight will probably not be at the center of media hype and attention as electronic money becomes more widespread. The ambition here is to contribute to that the most important advantages of electronic money are not lost in the process.

7.4 Parallel electronic currencies

A premise until now has been that a country absolutely needs to issue its own currency to have the necessary tools for macroeconomic control. So what can be done when the currency is issued by an institution above and outside the country? Examples are dollarised countries like Ecuador and El Salvador, and the eurozone countries. One of the hardest hit countries by the debt-induced crisis is Greece. In a number of papers I have since 2010 argued for the introduction of an electronic parallel (also called "complementary") national currency there (it would also work in other similarly crisis-hit countries). One paper is (Andresen; 2012), which is the basis for this section.

An indebted eurozone government has to extract euros out of the non-government economy to service its debt, by taxing more than it spends. The foreign-indebted private sector also extracts euros, sending these to creditors. The only way to (theoretically) counter these two "bloodletting" flows from a domestic economy is to increase net exports to a level that surpasses the sum of these two outgoing flows. This is exceedingly difficult, especially after debt service burdens have increased on the real economy, and because of idle production capacity due to the crisis. The other "way

out" is to sell off public property, which is unsustainable and economically destructive.

Debt could be partly written off and/or the debt service rates could be ameliorated, but to the degree the creditors refuse this, the domestic economy will be increasingly starved for money. Firms and individuals simply do not have enough of the instrument for the conducting of regular economic activity. This again leads to lower government income due to reduced tax payments and larger social outlays. The crisis is also amplified by increasing pessimism among individuals and firms: to the degree they possess euros, they hold back in spending, hiring and investment – and/or they move their money out of the country. All this contributes to further pessimism. We have an unstable downward spiral.

7.4.1 The proposal

Politically, both the EU elite and the elites in the crisis countries are strong supporters of the euro. There is also a majority in the general populace for sticking with the euro – mostly based on fear of what will happen if one reverts to a national currency. The mainstream advice seems to be to stick with the euro and hope for an internal devaluation of wages and prices to enhance the crisis country's competitiveness so much that future net exports will enable it to service its debts. This is a painful and slow process for the population (at best lasting several years, if working at all). Furthermore, the outcome is doubtful, especially since many trading partners are trying the same recipe.

A way out could be to furnish both households and firms with an additional domestic countrywide means of exchange, so that the large amount of unemployed may get jobs, and firms' spare capacity may be utilised. A euro-debt crisis country has a large output gap, and such a gap could be much diminished, without giving rise to significant inflation effects. Utilisation (and very fast activation) of this idle capacity may be achieved by nationally issued "electronic parallel money". This will quickly reduce unemployment and enable people and firms to exchange goods and services. It will also increase confidence, put a brake on the downward spiral, and even – as it will be argued below – enhance the circulation and net national acquisition of euros.

This proposed parallel medium of exchange will from now on be termed "emergency currency", abbreviated "EC". A unit of this currency will also be called "EC". The EC corresponds to HPM in one important and good sense: it cannot be lost.

7.4.2 How does it work?

Transactions are done (mostly) via mobile phone and automatically received and accounted for on a server with ample capacity at the country's Central Bank. Such a mobile banking system may be implemented through one of the technically proven schemes already in successful operation in some developing countries. There are no physical/paper EC's in circulation. The government (and local governments) have an EC account at the Central Bank. This account is debited whenever the government pays wages or pensions, or buys goods or services. All citizens and domestic firms have accounts there too, also interested foreign entities (but we will expect EC's to circulate only domestically in an initial phase). By this EC's are created ex nihilo, "printed" by the Central Bank.

The government pays employees, pensioners and suppliers both in EC's and euros. The proportions may be adjusted based on how the process develops. Taxes are also collected in a mix of the two currencies. The government-issued EC will be fiat money, and will have some intrinsic value since it may be used by the public to settle tax obligations (as argued by MMT). An EC will therefore be initially accepted to a fair degree as a means of payment by an agent – individual or firm – that is obliged to pay taxes.

Employees and firms offering goods and services will gradually – as the scheme gets more popular – decide to accept a certain share of EC's as payment, while the rest must still be in euros. While the government pays wages and taxes in a government-decided mix of the two currencies, the mix in private sector transactions may be decided freely by the involved parties, and will differ between trades. Both the government mix and private sector mix will necessarily have to be adjusted with time and circumstances. Employers and employees may locally negotiate the share of wages being paid in EC's, based on how things develop.

An additional positive effect of introducing EC's is this: By enabling activation of idle labour and production capacity, exports increase. Thus, even if this extra activity is mediated (partly) with EC's, this enhances the ability of the country to service its debt burden in euros.

Another positive effect is that pessimism is reduced. This will decrease the liquidity preference of individuals and firms that possess euros but have been holding back in their spending. Money velocity in euros will become greater: for a given amount of euro stock held by agents, the aggregate euro flow will increase, i.e. we get increased money velocity.

7.4.3 Discussion

The discussion will be done by addressing some counterarguments against, or expected questions about, the EC proposal.

The question of "confidence"

The EC is a *fiat* currency, not purely based on faith: A basic confidence is ensured because it may be used to pay (a share of) taxes, as already mentioned. One may in spite of this expect that initial confidence will be very low, not the least because of widespread distrust in authorities that until now haven't done much to ameliorate the effects of the crisis.

To discuss the prospects of an EC, it might be useful to define two entities, "*trust*" and "*need*". Even if trust is very low at the outset, need is very high: people and firms will have the choice of trying out an EC that is paid out to individuals and offered to firms, or let it accrue in their accounts or not accepting it in payment. Therefore some initial use of the EC should be expected because of the alternative of no work or no sale is considered even worse. Need will ensure some EC circulation, even if trust is low. With time, a positive feedback process will emerge: agents observe that transactions with EC's are happening, and this will increase trust, leading to more acceptance of EC's, and so on. This will – in addition to accepting EC's in payment – in the next round also encompass wages. When firms receive a share of EC's in payment, they will ask their employees to accept a share of EC's in their wages. And employees will often have the choice between accepting this, or unemployment. This again leads to firms becoming more willing to accept EC's in payment.

Inflation in EC's?

Assume that the government declares at the outset that the exchange rate EC to euro ought to be unity, and that firms are asked not to set prices in EC's high, but instead safeguard themselves in the start-up period by setting the initial EC *share* of an item's price low. What the government recommends will of course not necessarily be followed by vendors, but many will try this as a starting point. We should expect that firms (and individuals) that offer products or services where the dominant input factors are domestic, will be most willing to try offering a significant share of EC's in what they accept as payment. At the other end we have products that are imported, and the domestic input factors are subordinate: cars and petrol are examples. Here one can expect that only with time will such sellers

start accepting EC's, and the share will never become high. But there will be a mechanism at work in the right direction also there: when EC use has reached a reasonable and still growing level for other consumer items, for instance food (where domestic input factors are significant), import-based firms can negotiate a wage share being paid in EC's and the rest in euros, hence allowing also such firms to accept a share of EC's in the items they sell.

An important aside to this is that the existence of circulating EC's will enhance domestic output. To some degree this will lead to import substitution, *cet. par.* improving the balance of trade which is a good thing concerning the ability to service euro debt.

Regardless of possible government declarations, however, about how the parallel currency ought to be valued, one should expect the EC to lose value from parity with the euro. And floating the EC versus the euro must be accepted, there is no point in trying to uphold an artificially favourable exchange rate by this creating a black market. But the EC will be anchored not too far below the euro because one is allowed to pay a share of taxes with EC's, and one EC then counts as one euro. Also, as long as the economy is far away from full employment and firms have significant idle capacity, inflation pressures are not strong.

Euro debt and euro capital flight

One may at this stage correctly protest that introducing an EC does not in itself solve the euro debt problem. It also does not solve the problem of richer citizens moving their euros out of the country to avoid taxes or in fear of losses due to collapse of domestic banks.

To the first objection, one may reply that without a parallel medium of exchange an economy is wholly dependent on euros to uphold domestic activity. This puts the country in a very weak position when negotiating debt writedowns and/or lower interest rates and longer repayment times on existing debt.

Furthermore, by enabling the economy to run much closer to full capacity and employing a larger share of the population, the ability to export increases. The government receives more tax income, and can reduce its expenses for unemployed benefits and other social costs. The ability to service euro debt improves.

The problem of euro capital flight is not solved by introducing EC's, except that increased domestic economic confidence may after a while motivate many agents to repatriate their euros. But this will probably not make

a significant difference. Anyway, the issue of capital flight is there regardless of whether the EC proposal is implemented or not, and must be addressed somehow. And it has more serious effects without a parallel EC system in operation.

7.4.4 Summing up: far better than the bleak alternatives

A parallel electronic emergency currency will – with immediate effects – ameliorate the strongly and persistently lowered living standards for most people in crisis countries, which is the bleak and only future (lasting many years) that the EU and crisis country governments have been able to come up with. By the proposed scheme it should be possible to activate the immense underused potential that the hard-hit eurozone countries have, unemployed or underemployed people, to give many a better life and the country a return to social stability. It will primarily stimulate domestic production. It will also give euro-indebted countries a much better position in their bargaining for partial debt relief or less heavy euro debt service burdens.

Finally, it enables a gradual and controlled transition (back) to a national currency, if that is what is wanted. This proposal gives the national assembly in a crisis country the freedom to deliberate and make such a grave decision at any chosen time, and base it on experience with how the parallel currency and the economy have fared.

7.5 Conclusion on electronic money – and the rest

There are great possibilities for better control of macroeconomies with electronic money. The problem is not whether these would work – they obviously would. The problem is to get academic and public discussion, and – most important – implementations. There are working electronic money systems with an excellent track record on offer, see for instance (Economist; 2018),(Pandey; 2018). Doing this – for instance in Greece – is neither very expensive nor risky.

The challenge for the economics community – and the politicians that look to them for advice – is to leave behind the widespread unwillingness to think outside the box. Here J M Keynes (1936), ch. 12, hits the nail on the head:

"Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally."

I consider this an appropriate way to round off this thesis.

Bibliography

- Anderson, P. W. (1972). More is different, *Science* **177**(4047): 393–396.
- Andresen, T. (1998). The macroeconomy as a network of money-flow transfer functions, *Modeling, Identification and Control* **19**(4): 66–73.
- Andresen, T. (2003). Two feasible future scenarios: A high-tech utopia - and a high-tech dystopia, *Proceedings of the 5th Path to Full Employment Conference and the 10th National Conference on Unemployment*, University of Newcastle, Australia.
- Andresen, T. (2006). A critique of a Post Keynesian model of hoarding, and an alternative model, *Journal of Economic Behavior & Organization* **60**(2): 230–251.
- Andresen, T. (2012). A parallel emergency currency via the mobile phone network, *Die Parallelwährung: Optionen, Chancen, Risiken*, Bundesverband mittelständische Wirtschaft.
URL: <http://monneta.org/die-parallelwaehrung/>
- Andresen, T. (2013). Improved macroeconomic control with electronic money and modern monetary theory, *Real-World Economics Review* (63): 134–141.
URL: paecon.net/PAEReview/issue63/whole63.pdf
- Andresen, T. (2014). The central bank with an expanded role in a purely electronic monetary system, *Real-World Economics Review* (68).
URL: paecon.net/PAEReview/issue68/whole68.pdf
- Arestis, P. and Sawyer, M. (2007). *A Handbook of Alternative Monetary Economics*, Edward Elgar.
- Belangér, R. P. (1995). *Control Engineering: A Modern Approach*, Saunders College Publishing.

- Benes, J. and Kumhof, M. (2012). The Chicago Plan revisited, working paper.
URL: imf.org/external/pubs/ft/wp/2012/wp12202.pdf
- Bezemer, D. (2009). No one saw this coming. understanding financial crisis through accounting models, *Research Report 09002*, University of Groningen, Research Institute SOM (Systems, Organisations and Management).
URL: <http://EconPapers.repec.org/RePEc:gro:rugsom:09002>
- Bissell, C. (2007). Historical perspectives - the moniac a hydromechanical analog computer of the 1950s, *IEEE Control Systems Magazine* **27**(1).
- Blatt, J. M. (1983). *Dynamic economic systems : a post-Keynesian approach*, M.E. Sharpe, Armonk, N.Y.
- Bundesbank (2017). How money is created. website 25.04.2017.
- Casella, G. and Berger, R. (2002). *Statistical Inference*, Duxbury Press Thomson learning, Pacific Grove, CA.
- Dorf, R. C. and Bishop, R. H. (2017). *Modern Control Systems*, Pearson.
- Economist (2018). How mobile money is spreading.
- Eggertsson, G. B. and Krugman, P. (2012). Debt, deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach, *The Quarterly Journal of Economics* **127**(3): 1469–1513.
- Ferguson, B. and Lim, G. (1998). *Introduction to dynamic economic models*, Manchester University Press, Manchester, UK.
- Fisher, I. (1933a). The debt-deflation theory of great depressions, *Econometrica* **1**(4): 337–357.
- Fisher, I. (1933b). *Stamp Scrip*, Adelphi Company.
- Fisher, I. (1936). 100% money and the public debt, *Economic Forum* .
Spring Number, April June.
- Fontana, G. (2000). Post Keynesians and circuitists on money and uncertainty: an attempt at generality, *Journal of Post Keynesian Economics* **23**(1): 27–48.
- Gesell, S. (1958). *The Natural Economic Order*, Peter Owen Limited, London.

- Godley, W. and Cripps, F. (1983). *Macroeconomics*, Oxford University Press.
- Godley, W. and Lavoie, M. (2007). *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*, Palgrave Macmillan, New York.
- Goodwin, R. (1967). A growth cycle, pp. 54–58.
- Graziani, A. (1996). Money as purchasing power and money as a stock of wealth in keynesian economic thought, in G. Deleplace and E. J. Nell (eds), *Money in Motion - The Post Keynesian and Circulation Approaches*, Macmillan Press, Basingstoke and London, chapter B, pp. 139 – 153.
- Graziani, A. (2003). *The monetary theory of production*, Cambridge studies in international relations, Cambridge University Press, Cambridge, UK.
- Hicks, J. (1981). IS-LM: An explanation, *Journal of Post Keynesian Economics* **3**(2): 139–154.
- Huber, J. (2017). *Sovereign Money: Beyond Reserve Banking*, Palgrave Macmillan.
- Hudson, M. (2000). The mathematical economics of compound interest: a 4,000-year overview, *Journal of Economic Studies* **27**(4/5): 344–363.
- Hudson, M. (2010). Krugman, China and the role of finance, blog.
URL: <http://michael-hudson.com/2010/11/krugman-china-and-the-role-of-finance/>
- Hughes, N. and Lonie, S. (2007). M-Pesa: Mobile money for the "unbanked", turning cellphones into 24-hour tellers in Kenya, *Innovations: Technology, Governance, Globalization* **2**(1-2): 66–73.
URL: mitpressjournals.org/doi/10.1162/itgg.2007.2.1-2.63
- Keen, S. (1995). Finance and economic breakdown: Modeling Minsky's 'financial instability hypothesis.', *Journal of Post Keynesian Economics* **17**(3): 607–635.
- Keen, S. (2009). The dynamics of the monetary circuit, in S. Rossi and J.-F. Ponsot (eds), *The Political Economy of Monetary Circuits: Tradition and Change*, Palgrave Macmillan, London.

- Keen, S. (2010). Solving the paradox of monetary profits, *Economics: The Open-Access, Open Assessment E-Journal* 4(2010-31).
- Keen, S. (2011). *Debunking economics: The naked emperor dethroned?*, Zed Books, London.
- Keen, S. (2014). Oh my, Paul Krugman edition, blog.
URL: primeeconomics.org/articles/the-keen-krugman-debate
- Kennedy, M. (1991). *Geld ohne Zinsen und Inflation*, Goldmann, München.
- Keynes, J. M. (1936). *The general theory of employment, interest and money*, Macmillan, London.
- Kregel, J. (2012). Minsky and the narrow banking proposal: No solution for financial reform, *Public Policy Brief*, Levy Institute .
- Krugman, P. (2011). IS-LMentary, *New York Times* .
URL: krugman.blogs.nytimes.com/2011/10/09/is-lmentary/
- Krugman, P. (2012). Minsky and methodology (wonkish), *New York Times* .
URL: krugman.blogs.nytimes.com/2012/03/27/minksy-and-methodology-wonkish/
- Krugman, P. (2013). Debt in a time of zero, *New York Times* .
URL: krugman.blogs.nytimes.com/2013/01/02/debt-in-a-time-of-zero/
- Lavoie, M. (1992). *Foundations of Post-Keynesian Economic Analysis*, Edward Elgar Publishing.
- Lietaer, B. (2010). The Wörgl experiment: Austria (1932-1933).
URL: <http://www.lietaer.com/2010/03/the-worgl-experiment/>
- López G, J. and Assous, M. (2010). *Michal Kalecki*, Palgrave Macmillan, New York.
- Mankiw, N. G. (2008). *Principles of Macroeconomics, 5th Edition*, 5 edn, South-Western Cengage Learning.
- McLeay, M., Radia, A. and Thomas, R. (2014). Money creation in the modern economy, *Quarterly Bulletin*, Bank of England **Q1**.
- Minsky, H. P. (1992). The financial instability hypothesis. Economics Working Paper Archive, Levy Economics Institute.

- Nicolaysen, J. (2017). What should the future form of our money be? (speech held 25 April 2017).
URL: <http://www.norges-bank.no/en/published/speeches/2017/2017-04-25-dnva/>
- OECD (2017). Narrow money(M1).
URL: data.oecd.org/money/narrow-money-m1.htm
- Pandey, A. (2018). India's digital payment upstarts give visa, mastercard run for their money.
URL: <http://p.dw.com/p/2vYVd>
- Pettifor, A. (2013). The power to create money 'out of thin air' understanding capitalism's elastic production of money and moving on beyond Adam Smith and 'fractional reserve banking' - a review essay of Geoffrey Ingham's Capitalism, *Prime economics* .
URL: primeeconomics.org/articles/1443
- Phillips, A. W. (1954). Stabilisation policy in a closed economy, *The Economic Journal* **LXIV**.
- Phillips, A. W. (1957). Stabilisation policy and the time-forms of lagged responses, *The Economic Journal* **LXVII**.
- Piketty, T. and Goldhammer, A. (1995). *Capital in the Twenty-First Century*, The Belknap Press of Harvard University Press.
- Robinson, J. (1982). Shedding darkness, *Cambridge journal of Economics* **6**(3): 295–296.
- Rochon, L.-P. (2005). The existence of monetary profits within the monetary circuit, in G. Fontana and R. Realfonzo (eds), *Monetary Theory of Production: Tradition and Perspectives*, Palgrave Macmillan, Basingstoke, pp. 125–138.
- Samuelson, P. A. (1939). Interactions between the multiplier analysis and the principle of acceleration, *Review of Economic and Statistics* **21**(2): 75–78.
- Samuelson, P. A. (1948). *Economics*, McGraw-Hill, New York.
- Solow, R. (2008). The state of macroeconomics, *Journal of Economic Perspectives* **22**(1): 243–246.
URL: jstor.org/stable/27648233

- Sterman, J. D. (2000). *Business Dynamics - Systems Thinking and Modeling for a Complex World*, McGraw-Hill.
- Wray, L. R. (2012). *Modern Money Theory: A Primer on Macroeconomics for Sovereign Monetary Systems*, Palgrave Macmillan, New York.
- Zuckoff, M. (2005). *Ponzi's scheme: the true story of a financial legend*, Random House, New York.

Appendix A

Block diagram representation

This appendix explains block diagram symbolism. We will use the simple "households and firms" model in section 2.7 for explanatory examples. Some diagrams from there will be repeated below for convenience.

The rules for interpreting a block diagram are as follows:

1. The variable exiting a rectangular block is the product (also in a wider interpretation of the term, see point 4. below) of the variable entering the block and the coefficient or expression within the block. In figure A.1 we thus have $C = \frac{1}{T_H} M_H$. Several blocks along an arrow may be reduced to one equivalent block containing the product of the expression in each of the original blocks. More in point 4. below.
2. A circle at an intersection of lines signifies a *summation point*. The variable associated with an arrow leaving a circle is the sum of variables associated with arrows entering the circle. An arrow with a minus sign near it, means that the corresponding variable is to be subtracted in the summation. Thus we have $\dot{M}_H = W - C$.
3. A small dot upon a line signifies a *branching point*. This means that a variable is used as an input to more than one other part of the system. Example: we have the variable C being used both in $\dot{M}_H = W - C$ and $\dot{M}_F = C - W$.
4. The symbol " s " in the block diagram may be interpreted in two ways: as the free variable in the Laplace transformation (then all variables in the diagram are Laplace transforms, $C(s), W(s), M_H(s), M_F(s)$), or in the time domain as a differentiation operator (then we have

$C(t), W(t), M_H(t), M_F(t)$. We will elaborate on the second interpretation. The product-like construct " sx " is to be interpreted as $\frac{dx}{dt}$. Then " $\frac{1}{s}x$ " corresponds to $\int x(t)dt$. Expressions containing the s operator may be manipulated as if s was a regular multiplicative factor, along with constants. This means that $C = \frac{1}{T_H} \int \dot{M}_H(t)dt = \frac{1}{sT_H} \dot{M}_H = (\frac{1}{T_H} M_H)$.

5. If the block contains $\frac{1}{s}$ and a "time lag" like T_H and T_F in figure A.2 below, the expression in the block signifies a first-order linear differential equation. Consider the left part of figure A.1 We have $\dot{M}_H = W - C$ and $sT_H C = \dot{M}_H$. This may be combined to $sT_H C = W - C$, equivalent to the differential equation $T_H \frac{dC}{dt} = W - C$. We may now treat s as if it was a multiplicative constant, and solve $sT_H C = W - C$ for C , giving $C(t) = \frac{1}{1+T_H s} W(t)$. Here $W(t)$ is an input and $C(t)$ is the output. The expression $C(t) = \frac{1}{1+T_H s} W(t)$ corresponds exactly to $T_H \frac{dC}{dt} = W(t) - C(t)$. $\frac{1}{1+T_H s} = \frac{C}{W}(t)$ is called the *transfer function* from $W(t)$ to $C(t)$ ".

6. There are some useful (and easily derivable) rules for manipulating and/or simplifying block diagrams. Blocks along an arrow may be multiplied together to one equivalent block as mentioned above. Blocks on parallel arrows may be summed to one equivalent block on one arrow, and there is a rule for simplifying a feedback structure into a block without feedback, explained in figure A.3 below.

The block diagram of the "households and firms" model is shown in figure A.1.

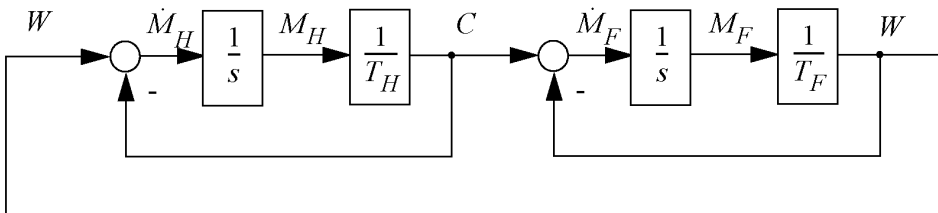


Figure A.1: Elementary block diagram of firms/households system

(Such a block diagram is called "elementary" when the blocks contain only constants or integrators.) If we reduce (simplify) the two inner loop

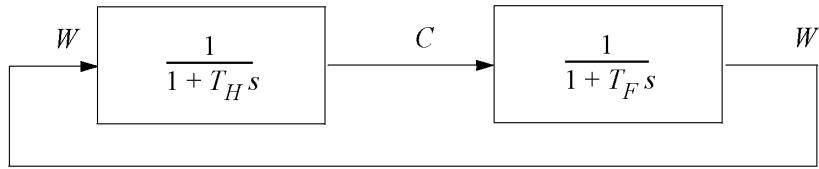


Figure A.2: Equivalent block diagram of firms/households system

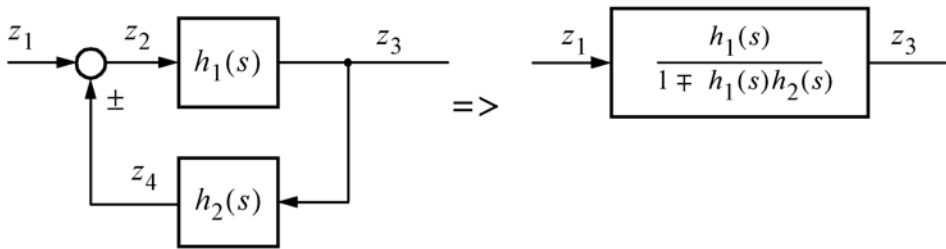


Figure A.3: Simplifying a feedback loop

subdiagrams, using the feedback rule explained in figure A.3, we get figure A.2.

Note that we reached the same result for the households subsystem by the derivation in point 5. above.

A similar block diagram symbolism applies to the time-discrete case. We refer to chapter 3. We use a time shift operator 'z' for discrete time. For a time-discrete variable $x[k]$ we define

$$x[k + 1] \triangleq zx[k] , \text{ and then } x[k - 1] = z^{-1}x[k] \tag{A.1}$$

This corresponds to a block diagram component as shown in figure A.4.

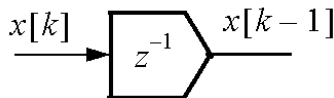


Figure A.4: discrete building block

The z^{-1} block plays a similar role in discrete-time systems as the integrator block $1/s$ in continuous systems. Transfer functions between any defined input and output become rational functions in z (or z^{-1} , the choice doesn't matter), just as continuous-time transfer functions are rational functions in s . Block diagrams for discrete time with the z operator can be manipulated and simplified with the same rules we use for block diagrams for continuous time. A discrete-time block diagram corresponds to a (set of) *difference* equation(s), while a continuous-time block diagram corresponds to a (set of) *differential* equation(s).

Appendix B

Biblical quotes against interest

"When your brother Israelite is reduced to poverty and cannot support himself in the community, you shall assist him as you would an alien and a stranger, and he shall live with you. You shall not charge him interest on a loan, either by deducting it in advance from the capital sum, or by adding it on repayment"
– Leviticus 25:35-36

"If you advance money to any poor man amongst my people, you shall not act like a money-lender: you must not exact interest in advance from him" – Exodus 22:25

"You shall not charge interest on anything you lend to a fellow- countryman, money or food or anything else on which interest can be charged. You may charge interest on a loan to foreigner but not on a loan to a fellow countryman..." – Deuteronomy 23:19-20

"O lord, who may lodge in thy tabernacle? The man who does not put his money out to usury" – Psalms 15

"He never lends either at discount or at interest. He shuns injustice and deals fairly between man and man" – Ezekiel 18:8-9

"..on the Day of Atonement, You shall send the ram's horn round. You shall send it through all the land to sound a blast, and so you shall hallow the fiftieth year and proclaim liberation

in the land for all its inhabitants. You shall make this your year of jubilee. Every man of you shall return to his patrimony, every man to his family.....In this year of the jubilee you shall return, every one of you, to his patrimony... if the man cannot afford to buy back the property, it shall remain in the hands of the purchaser till the year of the jubilee. It shall then revert to the original owner, and he shall return to his patrimony.... When your brother is reduced to poverty and sells himself to you, you shall not use him to work for you as a slave. His status shall be that of a hired man and a stranger lodging with you; he shall work for you until the year of the jubilee. He shall then leave your service, with his children, and go back to his family and to his ancestral property..." – Leviticus 25, excerpts

Appendix C

Corrected errors in the original submitted thesis

Corrections of typos and errors in the original thesis (but corrections of trivial spelling mistakes are not listed here):

Equation (2.20) was corrected 15-10-18, to:

$$\begin{aligned}\bar{T} = -h'(0) &= -(h_1 h_2 \cdots h_N)'(0) = -(h_1'(0) + h_2'(0) + \cdots + h_N'(0)) \\ &= T_1 + T_2 + \cdots + T_N\end{aligned}\tag{2.20}$$

p. 97, corrected 15-10-18: subscript B is removed from σ_B in figure 4.5.

p. 99, corrected 16-10-18: subscript B is removed from σ_B in figure 4.7.

p.127, figure 5.3, corrected 15-10-18 : "... and a CB" is removed in the figure caption

p. 138, eq. (6.5) shall be $\theta = \frac{D+D_K}{Y_o}$