Detzer, Daniel; Herr, Hansjörg

Working Paper
Theories of financial crises: An overview


Provided in Cooperation with:
Berlin Institute for International Political Economy (IPE)


This Version is available at:
http://hdl.handle.net/10419/92912

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Theories of financial crises – an overview

Authors: Daniel Detzer and Hansjörg Herr

Working Paper, No. 32/2014

Editors:
Sigrid Betzelt  Trevor Evans  Eckhard Hein  Hansjörg Herr
Martin Kronauer  Birgit Mahnkopf  Achim Truger
Theories of financial crises – an overview

Daniel Detzer, Hansjörg Herr
Institute for International Political Economy, Berlin School of Economics and Law (www.ipe-berlin.org)

Abstract: This paper analyses financial crises from a theoretical point of view. For this it reviews what different schools of economic thought have to say about financial crises. It examines first the approaches that regard financial crises as a disturbing factor of a generally stable real economy (Wicksell, Hayek, Schumpeter, Fisher, and the early Keynes). Thereafter, approaches, where the dichotomy between the monetary and the real sphere is lifted, are reviewed. Here in particular the later works of Keynes and the contributions of Minsky are of importance. Lastly, it is looked at the behavioural finance approaches. After having reviewed the different approaches, it is examined where those approaches have similarities and where they can be combined fruitfully. Based on this, we develop an own theoretical framework methodologically based on a Wicksellian cumulative process, however, overcoming the neoclassical dichotomy. The paper ends with some policy recommendations based on the developed theoretical framework.

Key words: Financial crisis, crisis theory, behavioral finance, Hayek, Keynes, Minsky, Schumpeter, Wicksell

Journal of Economic Literature classification: E03, E12, E13, G01,

Contact details:
Daniel Detzer, Daniel.Detzer@hwr-berlin.de; Hansjörg Herr, hansherr@hwr-berlin.de
Berlin School of Economics and Law, Badensche Str. 50 – 51, 10825 Berlin

Acknowledgments:
This paper is part of the results of the project Financialisation, Economy, Society and Sustainable Development (FESSUD). It has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 266800.
For helpful comments we would like to thank Nina Dodig, Trevor Evans, Eckhard Hein, Rainer Stachuletz, Natalia Budyldina, Barbara Schmitz, and the participants in the FESSUD Annual Conference 2013 in Amsterdam. Remaining errors are, of course, ours.

The views in this paper are the sole responsibility of the authors. The European Union is not liable for any use that may be made of the information contained therein.
1. Introduction

Financial crises are no new phenomenon. Even before market mechanisms dominated the whole economy financial crises were possible. Famous, for example, became the “Tulip Mania” in 1637 in the Netherlands which was a speculative bubble with extreme price increases of bulbs of at that time newly introduced tulips. When the bulb prices suddenly collapsed in a panic many speculators became over-indebted and became bankrupt. From the late 18th century on, when modern capitalism in England unfolded, financial crises were companions of capitalist development, however with different intensity in different historical periods. Already in the 19th century, economists started to develop models to try to understand financial crises. Karl Marx and John Steward Mill, among others, analysed financial crises. However, as the aim of this paper is not a history of economic thoughts we will not discuss those and other economists of that time as all their important arguments are included in later approaches about financial crises.

Knut Wicksell’s (1898) analysis seems to us a suitable framework to discuss financial crises. He delivered a framework of cumulative processes, and with it one of financial crises. This framework has been shaping economic thinking until today. Economists in the Keynesian and in the Neoclassical tradition followed his framework. Also, John Maynard Keynes and Hyman Minsky followed the Wicksellian approach, however, modified it in a fundamental way. Wicksell belongs to the Swedish school of neoclassical economists which in many ways stepped out of the traditional neoclassical model, which later became mainstream thinking. In the Wicksellian approach dynamic economic processes are explained by the interaction of two rates of return, which typically diverge or at least do not tend to equilibrium. Neoclassical and Keynesian economists agree that one of the two rates is the money interest rate. Which rate of return has to be compared with the money interest rate is solved differently in the two approaches. In the neoclassical paradigm it is the natural rate of return which is determined in the real sphere. In the Keynesian paradigm it is the expected rate of return for investment in the enterprise sector.

Three groups of models of financial crises can be distinguished.

The first group belongs to the neoclassical paradigm. All these models stress the point that the monetary sphere in one way or another can for a certain period of time develop independently of the real sphere. As, according to the neoclassical dichotomy and neoclassical thinking, the real sphere in the end dominates economic development, the monetary sphere sooner or later has to adjust. The general idea is that the monetary sphere in the short- and medium-term can become a disturbing factor for the real sphere. Irving Fisher and Milton Friedman, the icons of monetarist thinking in the 20th century, made this point very strongly. In extreme cases developments in the monetary sphere can lead to financial crises with far reaching repercussions for the real sphere. This group of models will be discussed in the second section.

John Maynard Keynes criticised the neoclassical dichotomy between a real and a monetary sphere. Instead of the dichotomy he proposed a model of a monetary production economy. In such an approach money plays a key role and penetrates all spheres of the economy. Instabilities and financial crises all develop within the framework of a monetary production economy. Keynes’ analyses in many respects covered
instability processes but he did not develop elaborated models of financial crises. It was Hyman Minsky who developed an explicit model of financial crises in the tradition of Keynesian thinking after World War II. Keynes and Minsky are presented in the third section of this paper.

There are economic approaches which concentrate their analyses on the isolated explanation of financial crises. These approaches almost exclusively concentrate their analysis on asset price inflations and the following asset price deflations and financial crises. In particular, behavioural finance contributed to this type of models, which will be discussed in section four.

In section five a synthesis of the different approaches is presented. We will look at the commonalities and differences of the approaches. Following this, we will develop our own idea of financial crisis on the basis of the Keynesian paradigm. Of course, not all approaches can be synthesised as the mixture of different paradigms would lead to inconsistent models. But some concepts from the approaches presented can be fruitfully used for an approach in a Keynesian tradition. For example, some ideas of behavioural finance can be easily integrated into a Keynesian-Minsky model. This section will also contain a sketch of international boom-bust cycles, where the analytical tools developed before can be transferred to the analysis of international crises.

In the last section of the paper conclusions are drawn.

2. Financial crises as disturbing factors of a stable real sphere

In this section neoclassical approaches to financial crises will be discussed. Here it is predominately focused on the old neoclassical economists. The reason for this is that modern versions of neoclassical thinking which base their models on the assumption of efficient financial markets and rational expectations are not suitable to explain financial crises. In these approaches an average economic agent acts on the basis of fundamentals, is perfectly informed about those fundamentals and all asset prices reflect fundamentals. Changes in fundamentals immediately lead to new equilibrium prices without any possibility of speculation. In this framework systemic crises simply have no place. The same is the case in General Equilibrium Models in the tradition of Léon Walras which cannot even introduce money to their models in any meaningful way.

We start the analysis with Knut Wicksell (1898, 1906). He explained the dynamics of capitalist economies by the interaction of two rates of return, the so-called natural rate of interest and the money interest rate. The natural rate of interest is the interest rate of the real neoclassical sphere in an equilibrium constellation. It is the interest rate which would be realized if the neoclassical capital market actually existed, savings and

1 In the Classical paradigm the “natural” price was the long-run equilibrium price in contrast to the market price which fluctuates around the natural price. In a similar way the natural interest rate has the meaning of an equilibrium interest rate determined by fundamentals.
investment would directly interact and there would be no money. “There is a certain rate of interest on loans which is neutral in respect to commodity prices, and tends neither to raise nor to lower them. This is necessarily the same as the rate of interest which would be determined by supply and demand if no use were made of money and all lending were effected in the form of real capital goods. It comes to much the same thing to describe it as the current value of the natural rate of interest on capital.” (Wicksell 1898: 102)
The money interest rate according to Wicksell is determined in the monetary sphere mainly by the central bank.

As soon as the money interest rate is lower than the natural interest rate a cumulative investment process is triggered. “The number of people becoming entrepreneurs will be abnormally increased.” (Wicksell 1898: 106) The resulting expansion leads to an inflationary process. Financing does not restrict expansionary processes as the monetary system is “elastic” (Wicksell 1898: 110). As soon as the money interest rate is above the natural interest rate a deflationary contraction process results. In Wicksell’s framework price level changes, and changes in real production and employment are not sharply distinguished. Both take place at the same time. Increases in output and prices in an expansion phase and shrinking output and prices in contraction phases are typical for all neoclassical approaches in the tradition of Wicksell.

It is important to point out that as soon as the two interest rates are not the same a cumulative process develops which has no endogenous tendency to tend towards equilibrium. For example, if the money rate is below the natural rate the economy will come into a situation of overheating which has no tendency to be corrected. Only the central bank can stop the inflationary expansion by increasing the interest rate. Money supply in this argumentation becomes endogenous whereas the central bank has to use the interest rate as the main policy tool. According to Wicksell, the natural rate of interest is not stable. Many changes in the real sphere can change the natural rate. Examples are changes in technology or changes of household’s preference which lead to a new saving behaviour. This implies that the central bank with its interest rate policy has to follow the natural interest rate in a discretionary way.

Wicksell’s approach serves as a suitable framework to explain financial crises. It stresses the character of capitalist development as a sequence of cumulative expansion and contractions which affect the whole economy. A cumulative expansion period is trigged exogenously. An expansion will lead to increasing instability and fragility and must sooner or later come to an end. It makes place for a cumulative contraction phase. A sharp enough contraction will lead to systemic problems in the financial system. The vision Wicksell develops is an economy which switches from cumulative expansion to cumulative contraction whereas the end of the cumulative processes cannot be explained endogenously in any strict sense.

Wicksell's model of cumulative inflationary and deflationary phases was taken over by a large number of economists in the early 20th century. Hayek in his business cycle models and Keynes in his “Treatise on Money” were especially important. For Hayek (1929, 1931) the modern credit organization involves a dilemma which is difficult to solve. Say’s law postulates that aggregate supply always exactly creates the aggregate demand to sell all products. For a barter economy this is obvious because in this case to offer a

---

2 Important were also Ludwig von Mises (1912) and Dennis Robertson (1940).
good also means to demand a good. A mismatch between demand and supply in a single market is possible, but there can never be a mismatch between aggregate supply and aggregate demand. Say’s law crucially depends on the non-existence of money hoarding. As soon as hoarding of money is possible a situation is possible where nobody wants to buy and everybody wants to sell. More important for Hayek is the interaction between savings and (net) investment. As soon as saving and investment exists it is not avoidable that credit relationships start to exist between households and enterprises. In the usual neoclassical capital market savings mean the supply of additional credit in a certain period and net investment means an additional demand for credit in the same period. The interest rate – an interest rate determined in the real sphere with the marginal productivity of capital on the one side and time preference of households on the other side – equalizes savings and investment in a way that Say’s law also holds in an economy with credit.

One of the key arguments of Hayek is that the modern credit system does not fit the vision of the neoclassical capital market. “… the influence of money should be sought in the fact that when the volume of money is elastic, there may exist a lack of rigidity in the relationship between saving and the creation of real capital.” (Hayek 1929: 101) Funds for investment flow from savings and money creation. Hayek was very much impressed by the German economist Albert Hahn (1920) who in a radical way pointed out that in modern credit systems, with commercial banks, business on the asset side of the balance sheet of a bank comes first and deposits follow. A credit, Hahn argues, creates a deposit and not the other way round. And the central bank has no direct efficient instrument to stop a credit and deposit expansion when banks want to give credit. Hayek basically agrees with Hahn. But now a problem is created. The credit volume given by banks does not necessarily reflect the volume of savings. Savings do not restrict the supply of credit any longer. In modern credit organizations the credit volume becomes independent of savings. If credit demand increases for investment, for whatever reason, the interest rate does not act as a “brake” for investment as it would if the neoclassical capital market would work as assumed according to the neoclassical model.

According to Hayek many factors can lead to an increase in investment without a corresponding increase in savings. As the credit system is “elastic” it will respond with credit expansion without bankers knowing that credit demand is not driven by long-term fundamentals. Competition in the banking sector supports such an endogenous credit expansion process. A bank which does not follow the general trend of credit expansion is flooded with liquidity in the form of deposits. This creates a stimulus for such a bank to give more credit as well. During such a process the monetary sphere becomes – if one likes – independent from the real sphere for a certain period of time. However, in the long run the real sphere will dominate the nominal sphere. During an expansion real investment materializes which in the medium-term is not profitable. An expansion not financed by savings leads to an “artificial” increase in investment which is not able to earn the required long-term rate of return which is given by marginal productivity in equilibrium. Sooner or later a credit expansion must come to an end. During the following contraction the artificially increased capital stock will be destroyed. A crisis with the potential of a financial crisis is the outcome.

John Maynard Keynes (1930) follows to a certain extent Wicksell and Hayek. However, he adds one important argument which later became one of the cornerstones of the Keynesian paradigm. From national accounting Keynes deducts his so-called fundamental equations of the value of money. In a closed economy the net domestic product, or national income, \((Y)\) equals wages \((W)\) plus normal profit \((Q_N)\) and extra or
windfall profits \((Q_E)\). Windfall profits are identical with undistributed profits of the enterprise sector. With \(P\) as the price level and \(Y\), as real income \((Y = Y_rP)\) it follows \(Y = Y_rP = W + Q_N + Q_E\). Isolating the price level the equation of the value of money becomes \(P = \frac{W}{Y_r} + \frac{Q_N}{Y_r} + \frac{Q_E}{Y_r}\). The term \(\frac{W}{Y_r}\) expresses unit-labour costs, the term \(\frac{Q_N}{Y_r}\) normal profits per unit of output or unit-profit costs.\(^3\) The latter represent interest costs but also “normal” dividends and other profits flowing to households. It is assumed that in equilibrium all profits are distributed to households and investment financed by profits is zero. Keynes, implicitly assuming monopolistic competition, argues that increases in unit-labour costs and profit costs are rolled over by firms and increase the price level.\(^4\) Falling costs, of course, lead to a falling price level. Cost and price level changes are independent of the demand constellation. A direct price-price effect is assumed as soon as the macroeconomic costs level changes. This means that in the framework of a closed economy and a comparative-static equilibrium analysis, firms are able to roll-over costs without excess demand. If all firms in an industry, for example, need oil as an input and are confronted with higher oil prices, they will increase their output prices. If nominal unit-labour costs in the industry increase, firms will react in the same way.\(^5\) This is an important difference to neoclassical models of price-level changes which only use excess demand or excess supply in the goods market as drivers for inflation and deflation. However, Keynes argues that increasing costs cannot be rolled over to prices in all constellations (see below).

In the long-run costs determine the price level. Thus, in the model presented above, costs depend on \(\frac{W}{Y_r}\) and \(\frac{Q_N}{Y_r}\). Normal profits per output unit can change, however not in the same dynamic way as nominal unit-labour costs. Unit-labour costs depend on wage per hour and labour productivity. This can be shown when nominator and denominator in \(\frac{W}{Y_r}\) are divided by labour input. Productivity changes at least in developed countries are not very high and relatively stable. This implies that the most important factor that determines changes in costs and the price level are changes in nominal wages.\(^6\) With this argument Keynes presented a key additional factor for cumulative expansion and contraction processes. For systemic financial crises deflationary developments in particular are of key importance. A falling price level directly increases the real debt burden of all debtors in domestic currency. The combination of goods market deflation and high debt irrevocably leads to a systemic financial crisis. This point was made very clear by Irving Fisher (1933) who developed his real debt theory of deflations under the impression of the Great Depression. Keynes did not discuss financial crises explicitly. But for him it was clear that falling nominal wages are a disaster. “Thus it is fortunate that the workers, though unconsciously, are instinctively more reasonable economists than the classical school, inasmuch as they resist reduction of money-wages.” (Keynes 1936: 14) According to Keynes nominal wages should not be flexible and nominal wages never should decrease – this is what

\(^3\) Normal profits are, as mentioned, considered as costs. Later in the General Theory Keynes argued that a liquidity premium would determine the interest rate and the interest rate the profit rate (Keynes 1930, chapter 17). Profit as a cost can be presented in a much more differentiated way (see for example Hein 2012 or Herr/Kazandziska 2011). But this is not the topic of this paper.

\(^4\) Of course, in oligopolistic and monopolistic markets firms also have power to change prices.

\(^5\) One consequence of this argument is that in a situation of unemployment higher nominal wages do not lead to additional real demand and not to an increase of output and employment up to full capacity utilisation.

\(^6\) Exchange rate movements can in a dramatic way change the domestic price level.
Keynes repeatedly recommended. Nominal wage developments should become a nominal anchor for the price level (see also Riese 1986, Herr 2009).7

Let us come to \((Q_E/Y)\). From national accounting it follows that windfall profits are identical with investment minus household savings. Thus we get \(Q_E = I - S_H\) with \(I\) as net investment and \(S_H\) as household savings.8 When net investment is higher than household saving there is excess demand in the goods market and the price level increases. In the opposite constellation a lack of demand leads to a goods market deflation. The more investment exceeds household savings the higher the increase in the price level and the higher undistributed profits. Windfall profits used for investment and consumption of capitalists or wealth owners lead to further demand and even higher profits, they become a ‘widow’s cruse’ which remains undepleted whatever is spend. On the other hand, the more firms cut investment expenditure in a situation of a loss the more losses they realise. Firms suffer from a ‘Danaid jar’ which never can be filled (Keynes 1930: 125).9 With the ‘widow’s cruse’ and the ‘Danaid jar’ Keynes introduced an additional dynamic element in his model which leads to cumulative inflationary and deflationary processes. These processes do not only have price level effects but are combined with real economic expansions and contractions. A clear separation between quantity- and price-effects does not exist. Only in the General Theory Keynes implicitly argues that in a situation of unused capacities higher demand leads to higher output. But higher demand in a situation of full capacity utilisation increases the price level. Of course there is an area of price-quantity effects when during an expansion period bottlenecks are reached in some industries and not in others.

Explicitly referring to Wicksell, the interest rate which balances net investment and household savings and makes undistributed profits \((Q_E)\) zero was called by Keynes the natural rate of interest. “Every departure of the market rate from the natural rate tends (..) to set up a disturbance of the price level by causing ... \((Q_E, \text{ the authors}) \text{ ... to depart from zero.}^7\) (Keynes 1930: 139) As in all Wicksellian frameworks money must be considered to be endogenous.

In the tradition of the Treatise on Money several constellations between cost inflation / cost deflation and demand inflation / demand deflation can be distinguished.

a) A combination of demand inflation and cost inflation leads to a cumulative expansion. A boom phase which, according to Keynes, can be caused by many factors typically leads sooner or later to such a constellation. An economic expansion leads to a demand inflation which leads to undistributed profits in the enterprise sector and typically stimulates further investment and a riotous living of profit receivers. The expansion will lead to higher employment and in the end higher nominal wages. Both inflationary forces now enforce each other and lead to a cumulative expansion.

7 Actually the nominal wage level should increase according to medium-term productivity development plus the target rate of inflation to become a nominal anchor for the desired inflation rate.

8 In national accounting \(Y = I + C\) with \(C\) as consumption demand. But it also holds \(Y = C + S_H + Q_E\) whereas \(Q_E\) are undistributed profits. Total saving obviously is \(S = S_H + Q_E\). From the two definitions of \(Y\) it follows: \(Q_E = I - S_H\).

9 The relationship between investment demand (including consumption out of profits) and profits was stressed by many economists (see especially Kalecki 1954).
b) A second cumulative constellation is a combination of demand deflation and cost deflation. A reduction of demand decreases output and prices and increases unemployment. Higher unemployment can lead to nominal wage cuts and a further stimulation of the deflationary process. As soon as the wage anchor breaks a cumulative deflationary wage price spiral together with a lack of demand and high losses of the enterprise sector develops. A systemic financial crisis is hardly avoidable then.

c) The combination of cost inflation and demand deflation characterises a constellation when, due to a lack of demand, firms are not able to roll-over higher costs. Such a stagflation leads to a profit squeeze. A stagflation is typical at the end of an expansion when central banks start to fight inflation. Of course, other explanations of a stagflation are possible, for example a collapse of investment as a result of a negative shock or price increase of natural resources in an overall stagnating economy.

d) Theoretically the combination of cost deflation and demand inflation is also possible. Such a constellation fits, for example, a demand stimulating policy in a country suffering from deflationary wage decrease.\(^\text{10}\)

Keynes in his General Theory criticised his own approach of the Treatise on Money including the concept of a natural rate of interest. Before we come to this point two further economists who contributed to crisis theories, Irving Fisher and Joseph Schumpeter, will be discussed.

*Irving Fisher (1911)* is the founder of the modern version of the quantity theory of money which was set out by David Hume and taken over by David Ricardo and almost all classical economists. Fisher argues, following the classical and neoclassical paradigm that in the long run money is neutral and changes of money supply in the end only affect the price level. But what is important here is that changes in the money supply can have deep and destabilising effects on the economy in the short and medium term. In the short term, money is anything else but neutral. Periods of “transition” from one equilibrium to another after an increase in the supply of money leads to “action and reaction” and “a cycle of ‘prosperity’ and ‘depression’” (Fisher 1911: 72). Later Milton Friedman (1968) argued in exactly the same way. Also for him, money can become a fundamentally disturbing factor for the real economy. His recommendation to follow strict monetary targeting as an economic policy rule had the purpose of enforcing the neutrality of money. For the older versions of the classical and neoclassical paradigm we can sum up: In spite of the hypothesis of the long-term neutrality of money the latter can become a fundamentally disturbing factor for the economy. Asset price bubbles and financial crisis are extreme versions of such destructive disturbances created in the monetary sphere.

This brings us to Irving Fisher’s (1933) famous debt-deflation theory. Under the impact of the Great Depression in the United States he wrote a paper which until today is one of the cornerstones to understand the destructive power of deflationary processes. For Fisher a business cycle with it ups and downs is difficult to avoid and also not a fundamental problem. A problem is, however, that under certain conditions a “normal”

\(^{10}\) Japan’s fiscal policy in the 1990s and 2000s fits this constellation.
business cycle can get out of control and a cumulative development is triggered which leads to a cumulative breakdown of the economy.

In an almost Keynesian manner Fisher argues that overoptimistic expectations lead to periods of expansion including asset price bubbles. Herding and speculative behaviour trigger asset price inflations, which are usually combined with huge credit expansion. These are also phases of high GDP growth and high employment. When the asset price inflation comes to an end an asset price deflation is the ultimate result. Asset price deflations are simply the consequence of previous asset price inflations. The bigger the bubble is the bigger will be the following asset price deflation. The end of an expansion period - as the beginning - depends on many factors and can only be clarified in a historical analysis. Asset price deflations lead to the destruction of wealth as well as to problems for speculators and other economic units in paying back their debt. Non-performing loans start to grow. Distress selling of assets to be able to service debt and panic leads to sharply falling asset prices.

Asset price inflations and following asset price deflations and non-performing loan problems are normal during capitalist development and over the business cycle. However, as mentioned, an economic downturn can get out of control and developments like the Great Depression in the 1930s can result. The difference between a normal crisis and a disaster leading to a systemic financial crisis with deep repercussions for production and employment is a goods market deflation. A constellation of high debt and goods market deflation leads to an increase of the real debt burden by all debtors in the domestic currency. The non-performing loan problem explodes, the coherence of financial markets erodes and the economic boat capsizes. “Then, the very effort of individuals to lessen their burden of debts increase it, because of the mass effect of the stampede to liquidate in swelling each dollar owed. Then we have the great paradox which, I submit, is the chief secret of most, if not all, great depressions: The more the debtors pay, the more they owe.” (Fisher 1933: 344) Fisher shows empirically that the deflation during the Great Depression increased the real debt burden in the US in spite of the fact that nominal debt had been paid back.

The key channel between an asset market deflation and a goods market deflation is the lack of demand for goods and services and at the same time an increase of supply because firms with non-performing loans try to sell everything in an attempt to survive. The lack of demand is also caused by collapsing investment and decreasing consumption demand. The latter decreases because of lower income and rising unemployment. Fisher remains a theorist of the quantity theory of money as the only channel of goods market deflation is a lack of demand in relation to supply. The role of the nominal wage level as a nominal anchor against deflation does not exist in Fisher’s thinking.

Nevertheless with his real debt effect of deflations Fisher stressed one of the most destructive and important effects which can create a financial crisis and even more important becomes one of the factors which leads to a cumulative breakdown of the financial system. What can be learned from Fisher is that a goods market deflation in a situation of high domestic debt is one of the worst things that can happen to a capitalist economy. And a capitalist economy without substantial debt is not imaginable. Minsky (1982: 393) correctly
wrote “when Fisher … identified the characteristics of a debt deflation process … [he] identified essential forces which make for the observed instability of capitalist economies.”

Joseph Schumpeter (1961) developed a model with a boom-bust cycle which can also lead to a financial crisis. Starting from an equilibrium situation some entrepreneurs start with an innovation (a new technology, a new product, a new organization, etc.). A stock of inventions is always available. It is the entrepreneur which selects some of them and triggers economic development. Entrepreneurship, which is very close to the Keynesian category of “animal spirits” (Keynes 1936: 161), plays the key role during an expansion process. According to Schumpeter, capitalist development cannot take place without credit. Credit is created ad hoc (out of nothing) (Schumpeter 1911: 107) by the banking system and gives the entrepreneur the financial power to get the physical inputs to implement the innovation. “The essential function of credit in our sense consists in enabling the entrepreneur to withdraw the producers’ goods which he needs from their previous employment, by exercising a demand for them, and thereby to force the economic system into new channels.” (Schumpeter 1911: 106)

Schumpeter then assumes a kind of herding behaviour of firms following the innovative entrepreneurs. The “followers” imitate the innovation to get some of the extra profits which can be earned in the new market. They are also forced to do so by competition. If they do not follow they will sooner or later be eliminated by the market. Driven by high investment and credit expansion a boom phase develops which at a certain point comes to end and gives way for a contraction. This expansion phase, where firms invest into the new innovation or take credit to reorganize and get more productive is accompanied by a second, often bigger and more visible phenomenon. Prosperity is accompanied by speculation. Companies speculate on the further expansion and increase orders and inventory. Speculation in the narrow sense may occur and lead to a bull market. Private households may take consumption loans, etc. The general expansion also leads companies to increase capacities (without increasing productivity) in anticipation of continuing high demand. Schumpeter divides credit into productive and unproductive credit, where the former increases productivity, while the latter does not. For Schumpeter the start of the expansion and its end depend on history and cannot be explained mechanically. However, he notes that as a precondition for a renewed expansion due to further innovations the occurred instability due to the previous innovation needs to settle. Additionally, during the introduction of the innovation, there is an increase in the price level. As soon as the innovative process comes to an end, the new equilibrium will be characterised by a lower price level. So the expansion is characterized by inflation, while the recession is characterized by deflation. What is important: during the crisis phase when demand and output shrink not all firms will survive. Firms which were not innovative enough will go bankrupt. In particular unproductive loans lead to problems. Economic crises and to a certain extent financial crises as well are part of the normal process of capitalist development and its capacity to increase productivity and innovate. Schumpeter (1942) speaks about a process of creative destruction which

---

11 See also (Vercelli 2011). The neoclassical model including the model with aggregate demand and aggregate supply argues that a deflation has stabilising effects as net monetary wealth in real terms will increase. This positive wealth effect then increases aggregate demand. Independent of the question whether net monetary wealth in a model with endogenous money exists such an effect must be considered as unimportant compared to the Fisher effect (see Tobin 1980).
is very close to Karl Marx’s (1867) idea of relative surplus value creation which is stimulated by the chance of extra profits innovative firms can earn. In Schumpeter’s theoretical model the economy falls back into an equilibrium constellation, however with a higher level of productivity and new products. A new expansion can start when entrepreneurial spirits are high again and finance is available.

However, Schumpeter (1961: 158f) also argued the crises can get out of control. In such a case not only less innovative firms are eliminated, also good firms break down. Such a development is dysfunctional and harmful for economic development. He makes a difference here between a recession, as the normal process that follows the expansion, and the depression, which describes this dysfunctional development. The depression is characterized by an undershooting of the new equilibrium, where cumulative processes (for example Fisher’s debt deflation) drive the economy further away from the equilibrium point. Whether the recession turns into a depression depends, according to Schumpeter, on external circumstances and is not predictable. However, some factors make it more likely to happen. Those are for example: the general mentality and mood in the business community and the general public, the extend of credit expansion (in particular unproductive credit) during the phase of prosperity, the extend of occurred maladjustment and of deceptive business practices.

3. The Keynesian approach with no dichotomy between a real and a monetary sphere

3.1. Keynes’ fundamental modification of the Wicksellian approach

In the General Theory Keynes criticised the Wicksellian approach especially on two points. Firstly, Keynes gave up the idea of a natural rate of interest. “In my Treatise on Money I defined what purported to be a unique rate of interest, which I called the natural rate of interest – namely the rate of interest which ... preserved equality between the rate of saving ... and the rate of investment. ... I had, however, overlooked the fact that in any given society there is, on this definition, a different natural rate of interest for each hypothetical level of employment.” (Keynes 1936: 242) With this insight the natural sphere as a reference point of the money interest rates does not exist any longer. Secondly, Keynes criticises the idea of all Wicksellians that credit supply is the sum of savings plus money creation by the banking system whereby the latter becomes a disturbing factor and creates all types of complications. This idea “has led to the worst muddles of all”. (Keynes 1936: 183) Wicksellians conclude that “if the quantity of money could only be kept constant in all circumstances, none of these complications would arise, since the evils supposed to result from the supposed excess of investment over savings proper would cease to be possible. But at this point we are in deep water.”

Keynes came to the conclusion that in a capitalist economy a market between savings and investment which is equalised by an interest rate simply does not exist. With this insight Say’s law also breaks down and the law of effective demand becomes the key to determining output and employment.

---

12 Keynes then quotes the Wild Duck by Henrik Ibsen which dives down to the bottom as deep as possible and is trapped in the weed and tangle and all the rubbish that is down there.
Instead of the concept of an interaction between a real sphere and a monetary sphere which can lead to instability and financial crises Keynes developed a model of a monetary production economy (Keynes 1933). He supported Karl Marx (1867) that the nucleus of a capitalist economy can be expressed by $M - C - (M+\Delta M)$ with $M$ as money, $C$ commodities and $\Delta M$ as profit. Money is invested in productive processes to earn more money. $C$ also stands here for the production process which is an income creating process as well. Both Keynes and Marx (1894) distinguish between management which is investing in productive capital and economic agents which finance firms. Marx (1894) symbolised a production process which is embedded in credit relationships with the formula $(M_{\text{Fin}} - M) - C - (M+\Delta M) - (M_{\text{Fin}}+\Delta M_{\text{Int}})$ with $(M_{\text{Fin}} - M)$ showing that management gets funds from financial institutions and private wealth owners $(M_{\text{Fin}})$ to invest in productive capital. From the money flowing back to firms loans have to be paid back to creditors including interest. This is shown by $(M+\Delta M) - (M_{\text{Fin}}+\Delta M_{\text{Int}})$. The important point is that credit given to firms and investment in productive capital must be considered as a potentially unstable process. For example, not enough or too much capital might be invested. Or the whole adventure may go wrong and end in a situation where credit cannot be paid back.

For Keynes uncertainty plays a central role in a monetary production economy. It is not only the uncertainty which exists in all societies, for example, about the harvest in the coming years; what is important is uncertainty created by the market mechanisms itself. Uncertainty means that not all future events are known and probability models are not able to overcome uncertainty. Economic agents know that there are known unknowns and even unknown unknowns. As the construct of a real sphere in Keynesian thinking does not exist, a real sphere also cannot serve as an anchor for expectations. It follows that expectations must be considered to become very unstable under certain conditions. Also economic agents can have different expectations. And expectations can be wrong. With hindsight former decisions may even look irrational, but at the time they were taken they might have been rational (Herr 2011). In the Keynesian paradigm for equilibrium models expectations have to be considered as exogenous. They depend on historical developments not only in the economic sphere but also with respect to social and political development and institutions. A lot of economic development then depends on exogenously given expectations. Under this perspective economic development can be modelled as a sequence of time periods all shaped from historical specific expectations. Long-term economic trends as a result of a development of a real sphere do not exist. Empirically measured trends are the result of the string of shorter sequences all shaped by expectations (Hahn 1981).\(^{13}\)

Uncertainty leads to certain techniques or behaviour by economic agents to cope with uncertainty. Keynes gives three examples of such a behaviour: “Agents assume that the present is a much more serviceable guide to the future than a candid examination of past experiences would show ... We assume that the existing state of opinion as expressed in prices and the character of existing output is based on a correct summing up of future prospects ... Knowing that our own individual judgement is worthless, we endeavour to

\(^{13}\) The school of Rational Expectations explains expectations endogenously. This is possible as it is assumed that expectations of (all or at least the average) economic agents is identical with the equilibrium outcome of the model. It is obvious that under such an assumption expectations cannot have a systematic effect on economic outcomes.
fall back on the judgement of the rest of the world ... The psychology of a society of individuals each of whom is endeavouring to copy the others lead to what we may strictly term a conventional judgment.” (Keynes 1937: 214) Many different techniques are possible to simplify decisions in a world of uncertainty and Keynes only analysed some of them. Especially the tendency towards a conventional judgment or “state of confidence” (as Keynes 1936, chapter 12, called it) leads to herding behaviour. Herding is one of the most powerful feedback mechanisms leading to unsustainable expansions and financial crises. Keynes is very close here to some ideas later developed within behavioural economics and especially behavioural finance.\footnote{This led Davidson (2010, 254) to call Keynes the “first behavioural economist”.} This allows the integration of research results from behavioural economics into Keynesian thinking (see for example King 2013).

Keynes (1936, 1937) followed Wicksell in modelling economic dynamic as the interaction of two rates of return. One of the two rates is the money interest rate. It is not simply fixed by the central bank; it is given by the intersection between the money supply function and the money demand function. Money supply is given exogenously by the central bank. From a modern Post-Keynesian perspective this is a step backwards compared to the Treatise on Money. But it allows Keynes to give credit supply and thus the creditor an active function in economic development.\footnote{In modern Post-Keynesian approaches an active role of credit supply is modelled in another way (see below).} Money demand depends on the demand for money for transaction purposes and hoarding. The latter fulfils the function to protect individual agents in an uncertain world from the imponderability of a capitalist economy and earns a liquidity premium. The marginal liquidity premium decreases with increases of money holdings for hoarding purposes. A household will hoard money as long as the marginal liquidity premium is higher than the money interest rates. This makes the demand for hoarding a function of the interest rate under the condition of a given level of uncertainty. The equilibrium interest rate is given at the intersection of money demand and money supply. An increase of the level of uncertainty leads to a higher level of the liquidity premium, a desire to hold more liquidity. The market outcome is an increase of the interest rate.

The second rate of return is the marginal efficiency of capital, a rate of return of investment expressed as an interest rate and calculated by management carrying out investment in productive capital. Keynes (1936: 135) defines it as following: “I define the marginal efficiency of capital as being equal to that rate of discount which would make the present value of the series of annuities given by the returns expected from the capital-asset during its life just equal to its supply price.” The marginal efficiency of capital has nothing to do with a marginal productivity of capital. It is determined mainly by the expected future yields an investment project creates, but also by the sum of money which has to be spent today to invest. The marginal efficiency of capital in the end depends on animal spirits or, as Schumpeter called it, entrepreneurship.

Keynes models economic dynamics via the interaction between the interest rate and the marginal efficiency of capital. As long as the latter is higher than the former investment will be carried out and an investment-income-creation process starts. As soon as the marginal efficiency of capital falls below the interest rate investment will collapse. The consequence is a fall in income and employment. As in Wicksell, there is no market mechanism which could adjust the two rates of return automatically or quickly. The opposite is the
case: When due to a negative expectation shock the marginal efficiency of capital collapses as future yields are calculated more cautiously at the same time the interest rate increases as demand for money for hoarding increases. In periods of optimism the marginal efficiency of capital increases and the interest rate at the same time decreases (see Keynes 1937, 1936: chapter 22). The interaction between the two rates leads to periods of high investment and high income creation and periods of low or no investment and shrinking income creation. Also possible in the Keynesian framework is a period of long-term stagnation, based on depressed animal spirits and a conventional judgement which does not change. There is no anchor for the marginal efficiency of capital in any kind of fundamentals as in a neoclassical tradition. Expectation may be anchored in institutions, but this is a different story. In Keynes’ approach the economy is part of society and there are many linkages between economic development and development in the society as a whole. Expectations are one of these linkages. This “openness” of the Keynesians system makes the economic system potentially very unstable. Financial crisis can be the result.

Keynes distinguished between the expectations of managers and expectation in the stock market. Managers like all economic agents can become more or less optimistic. However, their expectations are based on the situation in the industry and a certain level of information. In Keynes’ thinking agents in the stock market are less informed about the situation in an industry or do not even care about it. One consequence is that expectations in stock markets typically change more violently than expectations of managers. Stock markets can become a casino dominated by speculation. In the stock market even “professional investment may be linked to those newspaper competitions in which the competitor have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole” (Keynes 1936: 156). This kind of speculation can lead to self-reflexive developments in asset markets and to asset price inflations and bubbles which lead to fragile asset markets in general. Such developments are not restricted to stock markets, but to all kinds of asset markets like real estate, gold of currencies.

Let us come back to the topic of an endogenous money supply. Keynes (1936, 1937) gave creditors an active role in economic development. The disadvantage of Keynes’ approach here is that the money supply is given exogenously and the banking system is absolutely passive. The key role in determining the credit supply and the interest rate is taken over by wealth owners who increase or decrease their desire for hoarding (see above). Post-Keynesian models have since long followed the idea of an endogenous supply of money. The central bank sets the interest rate whereas the commercial banking system decides how much credit it takes. Credit expansion by commercial banks depends on credit demand by the enterprise sector or other borrowers. For macroeconomic credit supply wealth owners usually play a subordinate role as the commercial banking system together with the central bank dominates the credit market (for an overview of the broad literature about endogenous money supply see Lavoie 2011).16

In this type of theoretical setting there are two channels to give banks an active role. We start with the first one. Banks take credits to give credits. In addition banks are characterised by maturity transformation. That

16 For a model of endogenous money supply including the commercial banking system and deposit holding of the public see Heine/Herr (2013).
means they take short-term deposits or other forms of short-term debt and grant long-term credits. The interest spread between taking a credit and lending covers not only the costs of banking and bank profit. It also covers what can be called an uncertainty premium expressed in per cent. According to the level of uncertainty the uncertainty premium will be high or low. The existence of an uncertainty premium implies that the central bank cannot dictate long-term interest rates which are important for investment in productive capital. The power of the central bank is asymmetric. It can always push up the long-term interest rate by increasing the refinancing rate for commercial banks. But it cannot reduce the long-term interest rate to zero even if the refinancing rate is zero when banks have a high uncertainty premium.

More important than an uncertainty premium is credit rationing by the banking system. In a world of uncertainty not all credit demanders will be satisfied. It is one of the big short-comings in the standard model of asset markets to assume that credit demand and credit supply functions intersect and in this way an equilibrium can be found. Banks check whether a debtor has good collaterals, offers a good investment project, has high expected income, etc. Credit demanders who do not fulfil certain standards do usually not get a credit. Such credit demanders even do not get a credit when they offer high interest rates. A high interest rate cannot convince a creditor to give a credit when she or he expects that the credit cannot be paid back. In a situation of asymmetric information banks and other credit suppliers also will not increase the interest rate to the level the market would allow. They know that a high interest rate leads to adverse selection as the serious credit demanders leave the market and the risk-loving and desperate debtors with a high likelihood of default remain in the market. The intensity of credit rationing depends on the level of uncertainty. In a boom, when positive expectations prevail, banks are relatively lax to give out credit. When confidence breaks down in a crisis, banks will follow a policy of strict credit rationing. Credit rationing by commercial banks in a fundamental way reduces the power of central banks to stimulate the economy. In a situation of high uncertainty even very low interest rates do not help to overcome strict credit rationing.

From a New Keynesian point of view the case of credit rationing due to asymmetric information was demonstrated by Stiglitz and Weiss (1981) and Stiglitz and Greenwald (2003). Wolfson (1996) shows from a Post-Keynesian perspective that under uncertainty credit rationing can occur simply by asymmetric expectations, i.e. the same information is evaluated differently by borrower and lender.

Keynes never developed explicitly a model of financial crises. All elements for such a model exist but were not put together. Credit relationships and their development were never analysed by Keynes in a systematic way. It was Hyman Minsky who in the Keynesian tradition developed a Keynesian model with the explicit analysis of credit relationships and financial crises.

### 3.2. Minsky’s financial instability hypothesis

Since the 1960s Hyman Minsky developed his Financial Instability Hypothesis. Minsky’s most well-known works are his books “John Maynard Keynes” from 1975 and “Stabilizing an Unstable Economy” from 1986. He describes his contribution to economics as an interpretation of the essence of Keynes’ publication of “The difficulty to understand the functioning of credit markets is also created by the fact that in many cases the asset market is modelled as demand and supply of money. In this type of models the credit market is hidden behind the money market.
General Theory of Employment, Interest and Money” from 1936. His theory was also influenced by Schumpeter’s credit view of money. Minsky’s theory is based on two key theorems: firstly, an economy has financing regimes under which it is stable and financing regimes under which it is unstable, secondly, during periods of prosperity the economy can move from financing regimes which are stable to regimes which are unstable. Minsky, therefore, does not interpret financial instability as caused by exogenous shocks, but rather sees the occurrence of financial instability and financial crisis as inherent features of capitalist economies (Minsky 1992).

The theoretical basis
The following section will outline Minsky’s main theoretical concepts, which are: the distinction between hedge, speculative and ultra-speculative/Ponzi units, the characteristics of fragile and stable economic systems, the investment theory including the margins of safety, and the relation between profits and investment.

An economic unit has payment commitments out of its incurred debt liabilities. Those commitments are fixed in nominal terms and have to be paid at predefined dates. At the same time, it expects incoming cash flows to serve those commitments. Cash flows in turn are uncertain and depend on conditions in product markets, etc. Minsky defines income-debt relations for three types of economic units: hedge, speculative and ultra-speculative / Ponzi units.

A hedge unit has sufficient expected cash flows, so that it can serve all future payment commitments (interest and principal) when they occur. A hedge unit often exhibits a high share of equity finance. A speculative unit expects cash receipts that are sufficient to cover the interest payment, but not to pay back the principal. Only in the long-run, the cash flows expected are sufficient to reimburse the lenders in full. Until then, it needs to roll over its debt. Incoming cash flows of a Ponzi unit are not even expected to be sufficient to cover the interest payments in any period. Only during the final periods large cash flows are expected which will allow for repayment. Until then, it needs to capitalize interest payments on its balance sheet. Hence, it needs to roll-over the principal and find new financing for the accruing interest (Minsky 1982: 66-68, 105-108).

A financial system can be described as robust if small changes in cash flows, capitalization rates or in payment commitments will not inhibit the ability of most units to meet their financial commitments. The opposite is true for fragile systems. Here, small changes in the above mentioned variables will have detrimental effects so that many units fail to meet their financial commitments. Therefore, the economy becomes more fragile, when the relative weight of hedge units to speculative and Ponzi units declines. Because, in order to function normally a hedge unit only depends on the realization of its expectations regarding its cash receipts. Its payment commitments are determined over the entire period, so that changes in financial markets do not affect these. For speculative units, even if incoming cash flows are as expected, changes in financial market conditions can increase payment commitments due to changes in interest rates. Also, if speculative units are not able to roll over the principal after financial market conditions tightened, the unit may become troubled. Ponzi units depend even more on financial market conditions. Since they have to capitalize the accruing interest on their balance sheets, an increase in interest rates amplifies this process.
and the total amount of its payment commitments may soon exceed its incoming cash flows of the final period, so that the net-worth of a unit becomes negative and it cannot repay its debt (Minsky 1982: 66-68, 105-108).

Minsky’s theory of investment combines investment decisions of firms with their financing decisions and the willingness of lenders to provide external funds to them. This is important to determine the level of economic activity and also to explain the gradual move of the system towards instability.

Minsky modelled the investment decision as the interrelation between current output prices (for new investment) and asset prices (as the valuation of existing investment goods). Thus he does not follow Keynes, who used the marginal efficiency of capital and the money interest rate to explain investment. This is more than a technical question. Even though there are differences, Minsky seems to follow Tobin’s q, where stockholders’ and managers’ expectations and interests are identical. Current output prices of investment can be seen as the supply price for new investment. It depends on money wages, labour productivity, the short-run interest rate, and a profit mark-up. This price is regarded as stable in the short-run when capacity utilization is at normal levels. Prices in asset markets are the valuation of already existing investment goods and determine the demand price for new investment. The prices are determined by the present value of expected net cash flows, which depends on the state of uncertainty and the capitalization rate. For investment to take place, it is necessary that the demand price, determined in the asset markets, is above the supply price. Theoretically, if this condition is fulfilled and there are no financing constraints, firms would invest until the demand price of investment projects would be equal to their supply prices. Even if the demand price would not change, supply prices would start to increase as soon as full capacity utilization is realized. However, taking finance into consideration, investment demand is usually restricted before this point is reached. A firm will only be able to invest without changing its current debt and liquidity position as long as the investment is financed by free cash flows. The amount of investment higher than this requires external finance. If external debt finance is used, the borrower will incur fixed future payment commitments, while profits from the investment project remain uncertain. This increases the risk of insolvency when cash flows of the project are smaller than expected. Also, the risk for the lender that he will not be able to recover his loan in case the borrower becomes insolvent increases. This becomes more likely because, for example, the value of collateral in relation to debt decreases. Therefore, borrower’s risk and lender’s risk increase.18

Minsky modelled these risks in a way so that borrower’s risk decreases the demand price and lender’s risk increases the supply price. Thus, when firms keep increasing their borrowing, demand prices will progressively decrease and, at the same time, supply prices will continuously increase. As soon as both prices become equal, investment stops. In this circumstance investment will usually be much smaller than without the existence of borrower’s and lender’s risks. On the other hand, the discounts from the demand price and the surcharges to the supply price serve as margins of safety for the borrower and lender, and, therefore, increase the stability of the system (Minsky 1975: 93-116).

18 Already Keynes (1936) mentioned borrower’s and lender’s risk as additional factors for the investment decision. But Keynes did not analyze the relation in detail.
Following Keynes in Minsky’s approach investment is not determined by the physical productivity of capital goods. First, in an uncertain world the potential returns of an asset have to be evaluated by investors and depend on their personal optimism or pessimism about the future. In addition, the money rate of interest serves as the discounting rate. Given the interest rate, higher expected net cash flows lead to higher investment. Further, the uncertainty about future fulfils of debt commitments expressed in borrowers’ and lenders’ risks also influences investment. The larger those are, the lower will be investment activity (De Antoni 2006). Based on a behavioural argument Minsky introduced a feedback mechanism. If high margins of safety were chosen initially and expectations about future cash flows turned out to be correct, so that the incurred debt structures are validated, the borrower may realize that his initially chosen margins were too high and reduce them, which increases his investment demand. The same is true for the lender. With high margins of safety most borrowers will be able to meet their commitment, and, therefore, lender’s confidence about borrowers’ solvency increases. They will decrease their margins of safety and provide more funds, so that higher investment is possible. Based on the period of stability experienced, lenders and borrowers will then slowly start to revise their expectations and change their behaviour, which can endogenously lead to a boom (Minsky 1982: 71-99).

A crucial part of Minsky’s financial instability hypothesis is the Kaleckian relation between profits and investment expenditures. According to Toporowski (2007) this relation is the key argument to make financial fragility endogenous. The interdependence between the two variables allows for the cumulative processes of Minsky’s financial instability hypothesis. High investment and higher other demand elements lead to high profits, which in turn increase investment further. At the same time, high actual cash flows allow the fulfillment of all financial obligations. This way the income generating process is connected with the cash flows that are necessary for the validation of firms’ financing structures (Minsky 1982: 103-106). On the other hand, cash flows of the corporate sector plunge when investment begins to decline after the peak of the boom. Reduced cash flows in turn make it more difficult for firms to settle their financial commitments, which then can cause a crisis of over-indebtedness (Toporowski 2007).

**Minsky’s financial instability hypothesis**

Based on those theoretical foundations and the introduced feedback mechanisms, Minsky is now able to model how an economic system can create a boom and how it moves during the boom from financial stability to financial fragility, which then can turn into a bust.

**The boom phase – from financial stability to instability**

First, it is assumed that the economy has just gone through a crisis, whereby most speculative units have been wiped out. Correspondingly, borrowers and lenders are very conservative regarding finance after their recent crisis experience. Lenders demand high margins of safety. Borrowers evaluate profit perspectives of

---

19 In its simplest form under the assumption of a purely private domestic economy, where there is no savings out of wages and no consumption out of profit income from national accounting, it can be derived that profits are equal to net investment. If government spending, a foreign sector, consumption out of profits and savings out of wages are introduced it follows: after tax profits = gross investment + export surplus + government deficit + consumption of capitalist – savings of workers.
investment rather pessimistically and also demand high safety margins. This conservative debt structure leads to a phase of stability – investment and profits are low but stable. In this phase expectations are fulfilled, and incurred debt relationships are validated. According to Minsky’s behavioural assumptions the phase of stability leads to higher confidence. Borrowers get more confident in their profit expectations and are gradually willing to take lower margins of safety. Therefore, the effective demand price for investment goods increases. Stable profits and high margins of safety allowed lenders to register low default rates during this period, and their number of customers with a good credit history increased. Lenders, according to the assumption, use past credit histories to estimate the risk of potential borrowers in the future. Thus, they will reduce their margins of safety, i.e. they are willing to provide more finance at lower cost. Hence, the supply price of debt financed investment will be lower. A higher level of debt financed investment is now possible (Minsky 1986: 232-238).

There are further feedback mechanisms (see Figure 1). The higher investment leads to higher realized profits. Whereby, higher profits in turn have three feedback effects. Firstly, they allow borrowers to fulfil their past debt commitments. This validation of the incurred debt structure leads to higher confidence and encourages lenders and borrowers to reduce their margins of safety, and, therefore, increases the share of debt financed investment. Secondly, if the positive developments of current profits are regarded as permanent, it will increase expected funds for internally financed investment. Thirdly, it increases the expectations about future cash flows from new investment. These three mechanisms lead from higher initial investment, and therefore higher profits, to a further increase in internally and externally financed investment. This increased investment takes one back to the starting point in Figure 1, so that a first feedback loop is established, where increased investment leads to further increases in investment. De Antoni (2006) identified two further mechanisms that amplify the process. They are related to Minsky’s perspective on the determination of asset prices, the importance of liquidity and the process of endogenous money creation. Minsky held the view that, at given profit expectations and liquidity valuation, the quantity of money determines asset prices. If the quantity of money is high, asset prices will be high also, and if it is low, asset prices will be low. Additionally, following Keynes (1936, 1937) argument, the general public will have a greater desire to hold money when the future of the economy is highly uncertain. Holding money provides the holder with an insurance service against insolvency in the case that expectations about future cash flows are disappointed. The demand for money depends on the confidence people hold about the future (Minsky 1986). Minsky held an endogenous money view, which implies that money is created by banks’ credit expansions. If the debt financed part of investment, in Figure 1, is financed by bank loans, it will lead to an increase in the quantity of money, which, according to Minsky, will in turn lead to an increase in the demand for assets (with given portfolio preferences). Higher demand, subsequently, increases the market value of existing assets. The overall increased confidence, as a result of higher profits and the validation of expectations, will lower the desire to hold money. Therefore, the public will be willing to hold more illiquid assets instead. Respectively, a change in the portfolio allocation away from money towards non-monetary assets takes place, which again increases asset prices. Both mechanisms will increase the debt financed amount of investment in the economy even further (De Antoni 2006).
Figure 1: The self-amplifying investment loops of the Financial Instability Hypothesis

\[
\begin{align*}
I^{\uparrow} &\rightarrow \pi^{\uparrow} \\
\rightarrow Q_{i}^{\uparrow} \\
\text{Conf}^{\uparrow} &\rightarrow I^{\uparrow} = I_{i}^{\uparrow} + I_{\text{ind}}^{\uparrow} \\
\rightarrow M_{d}^{\downarrow} &\rightarrow P_{k}^{\uparrow} \\
\rightarrow M_{s}^{\uparrow} &\rightarrow P_{k}^{\uparrow} \\
\rightarrow I^{\uparrow} = I_{\text{ind}}^{\uparrow} \\
\end{align*}
\]

Source: own illustration, based on De Antoni, 2006

Notes: Conf = Confidence, I = Investment, \(I_{i}\) = Investment internally financed, \(I_{\text{ind}}\) = Investment debt financed, \(M_{s}\) = Money supply, \(M_{d}\) = Money demand, \(\pi\) = Profits, \(P_{k}\) = Asset prices, \(PV(\pi_{e})\) = Present value of expected profits, \(Q_{i}\) = Internal funds

During an investment boom the margins of safety are reduced and expected receipts exceed payment commitments only slightly and not throughout the whole project. Also, more short-term lending will take place. Therefore, the relative weight of Ponzi and speculative units increases during the boom. As a result, only small deviations from expectations can lead to defaults of borrowers. The system moves from a financial structure that is stable to one that is fragile (Minsky 1986: 244-245).

There is an often criticized inconsistency in Minsky’s theory. It is argued that investment undertaken by firms flows back towards them in the form of higher profits. This indicates that the debt-equity ratio does not need to increase during an expansion, and that his analysis suffers from a fallacy of composition (Lavoie and Seccareccia 2001). The inconsistency occurs because Minsky’s framework implicitly assumes a representative firm, suggesting that all firms in the non-financial corporate sector start investing at the same time. Under this condition each firm would, as a result of the investment, have higher profits and debt-equity ratios would not necessarily increase. However, as soon as many heterogeneous firms exist, the profits received from the investment of one firm will not necessarily correspond to that particular firm again. Therefore, while the debt-equity ratio of the sector as a whole may not deteriorate, some firms’ financial structures will (Toporowski 2007). Also, when workers, the government, or the foreign sector save, only a part of the investment expenditure will flow back as profits to the firms as a result.

The turning point, the bust and the crisis

In Minsky’s theory the boom is turned into a bust by an increase in interest rates. He argued the rising number of unfinished investment projects, which are still in the process of production, lead to an increasing share of inelastic demand for finance. As long as the supply curve for finance is infinitely elastic, this circumstance does not lead to an increase in the interest rate. According to Minsky (1982) this is only possible because the economy is either flooded with financial innovations, or the central bank is willing to supply an infinite amount of reserves. He questioned these possibilities and argued that either bottlenecks in the financial system will lead to an increasing interest rate, or inflationary pressure due to the boom will persuade the central bank to increase the interest rate (De Antoni 2006). Most economists, in the tradition of Wicksell and Minsky, can clearly explain why a boom can become continuously more fragile, which, in the end, must result in a collapse. However, when the economy collapses and how much fragility has been build-up can only be demonstrated by history.
When the interest rate increases and the downturn begins, the mechanisms described in Figure 1 work in reverse. The increase in the interest rate leads to a negative net-worth of some Ponzi units. Also, net worth of speculative and hedge units decrease. Credit ratings of borrowers are lowered. This together with the increase in the interest rate will raise the supply price of investment. Therefore, investment declines. Moreover, speculative and Ponzi financed firms get into trouble rolling over their debt, and can, therefore, not fulfil their commitments towards banks. Actual cash flows of banks fall below expected cash flows, so that these also need to raise new finance or sell assets. Furthermore, other units in the system will try to fulfill their commitments by selling assets. Therefore, asset values fall and the demand price for investment declines. Distress in the financial system and the shortage of liquidity further reduce investment. Due to this, profits decline. This decline in profits further depresses prices for assets. Additionally, the actual cash flows of firms are lower than expected. As a result, some hedge units become speculative units. Increasing defaults decrease the confidence of lenders and borrowers, who will realize their margins of safety were too low. Thus, they try to increase them. Borrowers’ and lenders’ risk increases and lead to an overall collapse of investment (Minsky 1982: 90-116).

If the demand price for investment collapses to such a degree that it is below the supply price, investment will not be undertaken. However, even if the demand price is still above the supply price, e.g. there are profitable investments opportunities, firms may not invest, because they have high debt ratios. The crisis shows them that their margins of safety are too low. Thus, cash flows will not be used for new investment projects but rather to reduce their debt ratio or to buy financial assets (Minsky 1975: 115-116). In this circumstance, firms’ main objectives are to repair their balance sheets. Purely financial considerations will negatively affect the demand for investment, and also influence the demand for consumption via the multiplier. Unemployment and depression are the result (Minsky 1986).

Whether this process turns into a deep depression or just a mild recession can be influenced by the action taken by the central bank and the government. Minsky argued that the main task of the central bank is to stabilize asset prices. This can either be done by direct intervention in the market or by providing liquidity to units that have trouble refinancing their positions. The government should increase its deficit during a downturn to stabilize the cash flows of firms. Depending on the actions of those two actors the end of the contraction is either on a higher or a lower level of economic activity than before (Minsky 1982: 110, 114).

Minsky’s work delivers, without a doubt, the most distinct explanation for the causes of financial instabilities and crisis in an economy. The basis of his theoretical approach is the Keynesian paradigm whereby he introduced the dimension of debt and debt-quotas into the analysis, something Keynes neglected. Minsky constructed a whole set of feedback mechanisms which lead to cumulative processes in the tradition of Wicksell. Many of them were later discussed in more detail in behavioural finance.

Charles Kindleberger explicitly built his historical analysis of financial crises on Minsky’s approach. He added, in some dimensions, to the understanding of financial crises (see Kindleberger / Aliber 2011).\footnote{The first edition of the book was published in 1978 with Kindleberger as the only author.}
Firstly, several factors that could lead to herding behaviour were distinguished, for example, the belief that others are better informed, the psychology of group thinking, a slow adjustment of individuals to a conventional judgement, hysteria in the light of increasing asset prices, and beliefs by investors in the wrong model (Kindleberger/ Aliber 2011: 42). Secondly, they made the argument that, in the last phase of a bubble, greed and fraud are common phenomena. Thirdly, they stress the role of the central bank as lender of last resort. Finally, their historical analyses forced them to take international crises and contagion mechanism explicitly into account.

4. Behavioural finance

4.1. Basics of behavioural finance

Behavioural finance tries to explain people’s economic decisions by combining findings of behavioural and cognitive research with traditional economics and finance. It was in particular driven by the failure of conventional theory (utility maximization of rational investors in efficient markets) to explain many empirical developments. Behavioural finance started developing in the 1960s and 1970s, when psychologists began to examine economic decision making processes. Among the most famous of those psychologists are Tversky and Kahneman, who detailed the heuristics and biases of humans that make decisions under uncertainty (see for example Kahnemmann and Tversky 1974). Their later developed Prospect Theory was awarded with a Nobel Prize. At the same time, increasing empirical findings that raised doubts about the validity of some key ideas in finance – the efficient market hypothesis and the capital asset pricing model – were collected. It was found that in many instances the behaviour and decisions of market participants did not fit the assumptions of standard theories. There is a vast and growing amount of literature in this field. Baker and Nofsinger (2010) identified four prevalent key themes: heuristics, framing, emotions and market impact.

Heuristics are means of reducing the cognitive resources necessary to find solutions to certain problems. They can be described as mental shortcuts and are often referred to as rules of thumb. They help individuals to make decisions under uncertainty with a limited ability to quantify the likelihood of outcomes. Researchers have identified a vast and still growing number of heuristics that have an effect on financial decision making processes. Among them are: affect, representativeness, availability, anchoring, adjustment, familiarity, overconfidence, status quo, loss and regret aversion, ambiguity aversion, conservatism, and mental accounting.

Framing means that the decisions taken by people, who have different options, are strongly influenced by how their choices are framed. That means people confronted with exactly the same problem, may chose different alternatives depending on how the problem is presented to them.

Emotions and associated human unconscious needs, fantasies and fears drive decisions. The underlying idea is that emotions and feelings influence psychic reality in manifold ways and areas. They may explain sudden changes in expectations and abrupt breakdowns of markets. Some of the research in this category investigates relations between investors’ moods that depend on factors like, sunshine, weather, or sporting
events and their investment decisions. Others found that most investors assume they are cleverer than the average investor and are convinced they can beat the market. It also was found that investors believe that an increase of asset prices would go on for ever – in spite of other historical fact.

The last key theme in this field of research focuses on market impact. This investigates whether and how cognitive errors and biases of individuals or groups affect market outcomes and prices and prevent financial markets to be efficient. In efficient markets agents use all information about the value of an asset available to them. In stock markets, so to speak, all agents calculate the future cash flows of all companies and calculate their present value. Under such assumptions asset prices would be ruled by fundamentals (Black 1986). Shleifer (2000: 24) stated that “market efficiency only emerges as an extreme special case unlikely to hold under plausible circumstances”. He named two key conditions in the theory that explain the deviations from efficient market outcomes – limits to arbitrage and investors’ sentiment. Limited arbitrage prevents the informed investors to trade away price deviations caused by irrational investors. Investors’ sentiment is responsible for disturbance of efficient prices in the first place. The research, for example, showed that if noise traders’ sentiment is unpredictable and correlated, arbitrage may not take place. If arbitrageurs have limited horizons and their risk bearing capacity is limited, they have to worry about having to liquidate the asset in mispriced markets. Therefore, the aggressiveness of arbitrageurs may be limited and large divergences between fundamental and market prices can occur (Shleifer 2000).

It is often argued that well informed arbitrageurs are investment funds managers who use outside investors’ money. However, the latter have limited information about fundamentals. Their investment decision is based on the past performance of funds. If there are noise traders who are responsible for the deviation of prices from their fundamental values, so the argument, even good informed and rational fund managers will follow the market trend and not fundamentals. If investment fund managers would try to gain through an arbitrage strategy and the prices in the short-term deviate still further from their fundamental value, funds would make a temporary loss. Due to the short-sightedness of outside investors this would lead to a withdrawal of funds away from the arbitrageurs, so that they may be forced to close their positions and realize their losses. By having to close the position they, then, also contribute additionally to the “irrational trend” and probably would lose their job as unsuccessful managers. As investment managers know these mechanisms, they do not follow arbitrage strategies (Shleifer 2000).

Another model presented by Shleifer (2000) becomes more explicit about the behaviour of noise traders. He showed that, in a case of positive feedbacks, even perfectly informed arbitrageurs will not follow an arbitrage strategy. There are different reasons for positive feedbacks. Studies found some investors base their investment decision on extrapolative expectations and trend chasing. Additionally, there are some technical features in financial markets that lead to positive feedback mechanisms like stop-loss orders, or liquidations after unfulfilled margin calls. Different from conventional theory, better informed arbitrageurs will not trade prices back to fundamental levels but will jump on the bandwagon and try to increase gains by anticipating

---

21 Noise traders are traders who react to irrelevant information (noise) or do not act rational according to the normative economic model (Shleifer 2000). Differently, information traders trade on the basis of all relevant information about the value of an asset (Black 1986).
the trading strategy of uninformed investors. Therefore, they will rationally amplify the bubble in the hope to sell assets before the bubble comes to its end. This case is very close to Keynes’ example of the beauty contest explained in the section about the Keynesian approach above.

Those examples of models from behavioural finance illustrate that financial markets do not work in the way the neoclassical model including the theory of efficient financial markets assume.

While there is a large catalogue of empirical findings of human behaviour and a growing volume of literature that tries to integrate these behavioural findings in models of financial markets, behavioural finance does not provide a unified framework or even an alternative paradigm to standard mainstream theory. Currently, there is a range of models which all focus on different specific points. However, there are models trying to integrate findings from behavioural finance with the aim to better understand asset price bubbles and why they are not contained by rational investors. In the following section, this paper will present the views of Robert Shiller, who combined many findings from psychology in a broader framework to explain the occurrence of bubbles.

4.2. Robert Shiller’s explanation of asset price bubbles

Many of the findings about human behaviour in financial markets were used by Shiller (2005 and 2012). Among other historical crises he looked at the stock market bubble around the turn of the twentieth century and at the real estate bubble developing in the US in 2005. Correspondingly, he derived a more general theory of bubbles. As a student Shiller was influenced by Charles Kindleberger, and, thus, similarities to Kindleberger’s perceptions of bubbles and financial crises are not surprising.

Shiller defined a speculative bubble as “a situation in which news of price increases spur investors’ enthusiasm, which spreads by psychological contagion from person to person, in the process amplifying stories that might justify the price increases and bringing in a larger and larger class of investors, who despite doubts about the real value of an investment, are drawn to it partly through envy of others’ successes and partly through a gambler’s excitement.” (Shiller 2012: 245)

He described the development of a bubble as follows:

1) There are some initial price increases (precipitating factors).
2) These are supported by certain feedback loops, which lead to further price increases.
3) The price increases draw the attention of the media and the general public towards some stories, which explain the increases. Here, particularly new era stories are of relevance. A new era story implies the general perception that the future is brighter and less uncertain than it was in the past.
4) Those stories justify the price increases.
5) This draws more people into the market to buy, which in turn increases prices even more.

The precipitating factors are exogenous, unique and historical developments, which lead to an initial upward movement in prices, that can, however, not explain the bubble alone. To explain how the initial price increases reproduce themselves, Shiller used certain feedback loops. Here, he mentioned direct price-price
feedbacks, where further price increases are triggered via investors’ enthusiasm and expectations based on initially increasing prices. Additionally, he named feedback mechanisms that are propagated through the real economy. In this regard, he argued there are effects where increasing asset prices lead through a wealth effect to higher economic activity and company profits, which in turn lead to further increases in demand for assets and prices. He gave a particular role to the news media, which shapes public opinion and categories of thought, and actively directs attention of the audience to certain stories. By doing this news media can strengthen feedback mechanisms. Also, they can be responsible for attention cascades, where the attention of the audience is directed to a certain fact that, then, leads attention to a range of other factors leading to changes in public perception. New era stories also play an essential part in the creation of a boom. The stories justify the enormous price increases. However, the stories do normally not cause the boom, but rather emerge as a subsequent interpretation of an asset market boom that has advanced already for a while (Shiller 2005). In each loop the new era stories spread further and bring new demand to the market, which leads to price increases. For further price increases, the new era stories need to spread so that more people are drawn into the market. Therefore, with each loop the stories become more prevalent and the price increases further. Shiller did not propose an endogenous turnaround that leads to the end of a bubble. There are countervailing forces from those that know what is going on and are willing to bet against the bubble. However, those activities are limited, since it is hard to determine with precision whether there is a bubble or not and it is even harder to determine its turning points (Shiller 2012).

Shiller (2005: 147-148) tried to advance from popular theories of bubbles, which talk about euphoric or panic-stricken behaviour where investors blindly follow the herd. He tried to give his theory a better foundation by reverting to findings from social psychology to explain the behaviour of investors. There is a range of what he calls “brain bugs” active, which influence peoples decision making. Speculative bubbles are the result of those brain bugs affecting the entire financial system (Shiller 2012). In the following section the most relevant of those psychological factors will be outlined.

There are patterns of human behaviour that suggest anchors in the market that would not persist if markets were entirely rational. Quantitative anchors indicate the appropriate market level and people use them to determine whether the market is over- or under-priced. Consequently, those anchors give them an indication for their buying and selling decisions. Here, he drew conclusions from a psychological phenomenon. If people have to answer a quantitative question about a number and the situation is ambiguous, they take whatever number is at hand as an orientation. He assumed this can also be applied to markets, so that if people are supposed to make judgments about the level of stock prices, the most likely anchor is the most recently remembered price. This would explain certain continuity in prices. The anchor can also be the rate of change. This means when prices went up by a certain degree in the recent past, people, not sure about future changes, would simply extrapolate past growth rates (Shiller 2005: 148 - 149).

Moral anchors are psychological factors that influence the decision of people about holding their investment or alternatively selling it and consuming the proceeds. When asset prices increase, people holding the asset will become wealthy. When prices become high enough, a discrepancy between the consumption and the wealth level of people will occur. People will be induced to sell part of their assets to increase their consumption to a higher level. This would moderate the asset price increases. Therefore, for the stock prices
to stay on a high level or to even climb further, from a certain point on people need to have strong reasons to
not sell their assets. Those moral anchors are not quantitative but take the form of storytelling and
justification. One example of such a story he gave was the narrative spread in the 1990s, that most people
did not become millionaires by earning exceptional incomes but rather by being frugal savers. According to
him, this was partly the moral anchor needed for sustaining the bull market in the 2000s (Shiller 2005: 149-
152).

Shiller (2005: 152-155) described some psychological findings about human nature that he also saw related
to the behaviour of asset markets. According to those findings, people are overconfident. Indicating that
people assume their judgment or reasoning is right with a high probability. This overconfidence strengthens
psychological anchors, because people tend to believe strongly in stories or reasons when they have
adopted them initially. Another related finding is the one of magical thinking. People believe that if they take
certain actions, they will be lucky, even though they should know by rational thinking that their action will not
affect the probabilities of the outcome. This can translate, at least at some intuitive level, in thinking of the
form “if I buy this stock, it will go up”, and therefore contribute to overconfidence and speculative bubbles.
Shiller also named representative heuristics which is basically judging the future outcome from observed
past outcomes. This strengthens the quantitative anchors. Overall, market participants seem, according to
questionnaires by Shiller, to make decisions by intuition and due to their overconfidence put great trust in this
intuition and the stability of anchors.

However, the anchors are fragile, which can be explained by non-consequentialist reasoning. This reasoning
is characterized by the inability to think about the elementary conclusions one would draw if certain
hypothetical events in the future were to occur. According to this finding, humans cannot make certain
decisions in advance, since there are emotional elements that one cannot appreciate before the event
actually occurs. Therefore, only when the event occurs, can one discover how one “feels” about it and then
make the decision based on this newly discovered information. Therefore, it is hard to predict the breaking
of a psychological anchor, because people discover this information only after the event or the price movement
has occurred (Shiller 2005: 155-156).

While these anchors work on an individual level, it is hard to suggest that all people at the same time suffer
from the same faulty thinking. Therefore, Shiller (2005: 157-159) again used findings from psychology that
could explain the synchronizing of thought and behaviour among investors. He argued that people are ready
to believe majority opinions or the opinions of an authority even if those plainly contradict matter of fact
judgments. Therefore, if a critical mass is achieved in a market believing in a certain story, it is likely that
more people will convert to this story and act accordingly.

Herd behaviour can also be caused by information cascades. It is assumed that economic agents have only
certain information about a situation. Decisions by individuals are taken in sequence, so that the second
decision taker can observe the decision of the first one, the third decision taker can observe the decision of
the first and the second one, etc. However, they do not know the reasons for those decisions. Now, while
they all have the same fundamental information, the second decision maker has the additional information
about the decision of the first decision maker. If the fundamental information is ambiguous, it may be rational
for the second decision maker to assume that the first decision maker had some additional information, which led to his choice. Therefore, he may make the same decision. Accordingly, when the information the actors possess is not revealed, herd behaviour can become rational. Shiller (2005: 160) gave an example for this. Take the decision of choosing between two restaurants, where the only information provided is the appearance of the restaurant from the outside. The first visitor only has this information for his decision, while a second visitor can see the first decision maker from the outside and may assume that the first decision maker knows that the chosen restaurant is good. Thus, the second visitor will follow the first one. The third decision maker sees now two persons in the restaurant and taking this as a quality signal follows the first visitor as well. Therefore, a situation can occur, where all restaurant visitors eat in the same restaurant, while the other restaurant is empty, without there being any objective difference between them.

There are additional reasons for exuberance and herd-like behaviour in the way humans communicate. A first point is that humans are more likely to communicate stories about interesting things. For example, if it is about stock markets it is rather the story about a hot stock, than the plain conversation about price-earnings ratios. Therefore, usually there takes place an exchange of stories rather than of fundamental information. Another interesting finding is that humans can have conflicting ideas coexisting in their minds. While they can think markets are not predictable, they may at the same time think they can make market forecasts. One explanation for this phenomenon is that very often ideas were received from some expert and were believed. Humans assume, despite the fact that two opinions contradict each other at a first glance, both ideas are correct and experts have on a professional level solved the perceived contradiction. Depending on the outside circumstances one or the other view may become dominant, which may explain sudden fluctuations and swings in markets (Shiller 2005: 160 – 170).

Social based selectivity of attention is another human characteristic. Humans, normally, can only pay attention to a certain small number of things at the same time. Additionally, they tend to focus their attention on things others focus their attention on. This can lead to shifts of focus on certain events or information and to the disregard of other information collectively and so lead to herd behaviour in the markets (Shiller 2005: 170).

Behavioural finance shows that investors do not act in a rational way as implied by rational expectations and the efficient market hypothesis. It tries to give a more accurate picture of human behaviour in financial markets. Additionally, it shows that even though there might be rational well informed investors in the market, their ability to correct prices is limited. Behavioural finance shows that cumulative processes are part of market processes in financial markets, and bubbles in financial markets are possible. Shiller tried to combine different findings from the field of behavioural finance to create a theory of asset price bubbles. He emphasized the importance of external factors to start the bubble, which then are amplified by feedback loops. Unique to Shiller's explanation of bubbles is the importance he gave to stories necessary to sustain the boom. He underpinned his theory with findings from other sciences such as psychology and sociology, which explain herd-like behaviours. These observations made the theory particularly valuable. Even though, his writing has many loose ends and undrawn conclusions it can fill in unexplained gaps and lose ends in other theories presented here. Without doubt Shiller gave a deeper understanding of Keynes (1936 and 1937) argument of conventional judgement or the state of confidence.
One consequence of behavioural finance is that asset prices do not reflect fundamentals and only by change reach fundamentals. Also in the long-term trend development of market prices does not follow trend development of fundamentals. Andrei Shleifer makes this clear. Each explanation drawn from behavioural finance and behavioural economics in general “has some intuitive appeal, and, each may account for some piece of the puzzle. Each, moreover, has some support from the lap or from other market data. Yet we still do not know which one of them is driving U.S. stock prices today, or which will drive them in the future. In fact, we do not even know that we have the right theory on the list” (Shleifer 2000: 185-186). However, the role of fundamentals in behavioural finance remains ambiguous. It is argued that economic agents do not act according to neoclassical assumption and even economic agents acting according to neoclassical assumptions do not determine prices as arbitrage processes are limited. This leads to the question which role fundamentals play when they are unimportant for markets. One could argue that clever economists can calculate fundamentals. But there is no guarantee that clever economists do not follow one of the behaviours described by behavioural finance. It seems to be theoretically much clearer to assume that there are no fundamentals which determine asset prices. In a world of uncertainty future cash flows simply cannot be objectively calculated. In such a Keynesian world expectations are given exogenously and asset prices have no anchor in fundamentals whichever.

5. A Synthesis of theories of financial crises

In this section three points will be discussed. First, the different financial crises theories are compared. Second, we present our own model of financial crises. Third, it is discusses what can be learned for international crises which are not in the centre of this theoretical debate.

5.1. Comparison of financial crises theories

In spite of deep differences between various paradigms, there are several important points all financial crisis theories share.

Firstly, in all models financial crises are the result of an unsustainable boom phase. The boom phase in asset markets can create asset price inflations, or it can, in the sphere of production, create overcapacities via too high investment and inflationary developments. Additionally, both spheres can be affected by boom phases. For example, Wicksell and Hayek stress more booms in the sphere of production, whereas behavioural finance concentrates on processes in asset markets, and Fisher, Keynes and Minsky integrate both spheres in their models.

Secondly, in all financial crises models feedback mechanism play the key role during expansion phases as well as during contraction phases. There are objective and subjective feedback mechanisms. Objective feedback mechanisms are, for example, developments in asset prices and product market prices, wealth effects, changes in income, real interest rates effects, effects on cash flows, etc. which improve or worsen the economic situation of an economic unit and lead to certain reactions. Subjective feedback mechanisms
are, for example, changes in perceptions which lead to more positive or more negative expectations, and, in this way, change the behaviour of economic units. Here mass psychological phenomena are discussed which can among many other things lead to exuberance and panic. Different financial crisis models stress different feedback mechanism. In neoclassical financial crisis models subjective feedback mechanism are in the background whereas subjective feedback mechanisms are stressed by behavioural finance. Financial crisis models in the Keynesian tradition stress both feedback mechanisms.

Thirdly, in all models of financial crises exogenous factors trigger an expansion or a boom. It can be a new innovation, a deregulation of a certain area in the economy, the end of a war, the election of a political party, a period of unjustified low interest rates, to name only a few. In Wicksell and Hayek as well as in the Keynes of the Treatise on Money as soon as the money interest rate is below the natural interest rate a cumulative expansion is triggered. However, these economists argued it would be wrong to hold central banks responsible for boom-and bust phases. According to them many factors can change the natural interest rate. The money interest rate has to follow the natural interest rate in a discretionary way. As monetary authorities cannot observe the natural interest rate, this can become very demanding for monetary policy makers and requires a high level of theoretical and historical understanding. In Keynes book “The General Theory of Employment, Interest and Money” and Minsky’s theory exists a comparable dynamic. However, they substituted the natural interest for the expected rate of return for investment. During the boom phase and before the crisis a cumulative development towards more and more fragility takes place in all models. In the end it is a question of taste to assume an endogenous or exogenous end of a boom. It must come to an end! However, which factor will stop the boom and when is open to history.

Fourthly, all approaches assume implicitly or explicitly an elastic financial system which is able to create sufficient endogenous credit to feed expansion. In some of the approaches, for example, in Wicksell’s, it is clearly stated that the central bank has to follow a certain interest rate policy and credit expansion and money creation is endogenously given by markets. The elasticity of the financial system can also be created by other theoretical assumptions. Keynes used in his book “The General Theory of Employment, Money and Interest” and in some of the papers that followed hoarding and dis-hoarding as well as the finance motive to give the economic model the elasticity to cumulative expansions and contractions. Keynes approach is, at this point, unsatisfactory as there is no clear argument for an endogenous money supply. It is not convincing that hoarding and dis-hoarding can substitute an endogenous approach of money creation. Minsky and many economists in the Post-Keynesian camp remedied this shortcoming in Keynes’ original approach.

In spite of the common points mentioned above, there are fundamental differences between the different approaches. Firstly, from a Keynesian perspective the interaction between a natural rate of interest and the

22 After World War II for neoclassical economists the central bank itself became the most important disturbing factor as soon as it follows discretionary monetary policy. On the one hand central bank may not be able to understand economic development (for example the time lack between an interest rate change and it effect on the economy), on the other hand politicians may force central banks to short-sighted policies to achieve short-term positive economic effects (for example before an election). Monetary targeting and independent central banks therefore became the policy recommendations for these types of monetarists (see as a representative Friedman 1969).
money interest rate is misleading. There is no natural rate of interest which depends in equilibrium on the marginal productivity of physical capital. To cut the economy in a real sphere, where money does not play a role, and a monetary sphere is a theoretical construct which is not useful to understand a capitalist economy. The natural rate of interest will change according to developments in the real sphere. Such changes can lead to disturbances in the economy and to financial crises. However, it must be assumed that disturbances in the Keynesian approach of the General Theory are potentially much deeper including the possibility of long-term stagnation. The marginal efficiency of capital, as Keynes called the expected rate of return of investment, has no anchor in the real sphere. It is driven by expectations based on a state of confidence or conventional judgment and can change in a quick and violent way triggering expansions and contractions. State of confidence and conventional judgment are reflecting the institutional, social, and political constellation in a country and connect the economy with developments in society.

Secondly, behaviour finance does not present a comprehensive macroeconomic approach. It focuses on asset market developments and sharpens the understanding of herding and other phenomena in asset markets. This is not a critique, but it gives behaviour finance a specific status. Behavioural finance can and is used in different ways. It can be used to show that the neoclassical approaches of efficient financial market hypothesis and rational expectations are not realistic. Behavioural finance makes clear that there are no fundamentals as an anchor for asset prices. Different behaviours lead to different market equilibriums and it is not possible to predict which behaviour will dominate at what time and whether a new behaviour can dominate markets at any time. Behavioural finance does not belong to a certain economic paradigm. It can be incorporated in different paradigm. Adding conclusions from behavioural finance in neoclassical models allows to show a large bundle of "exceptions" from the neoclassical equilibrium. Most work in the field of behavioural finance seems to be in such a tradition, and it does not try to become part of a new alternative paradigm fundamentally criticising neoclassical thinking. But behavioural finance can also be used to deepen the understanding of the economic system in the Keynesian paradigm. Keynes developed ideas which were very close to those that were later developed in behavioural finance. This allows the integration of research results from behavioural economics into Keynesian thinking.

Thirdly, from a Keynesian perspective Minsky delivered the most comprehensive model of financial crises. His framework, directly based on Keynes but adding explicitly indebtedness of economic units, can become the basis for further developing models of financial crises and their interrelation with the rest of the economy.

5.2. A model of financial crises
To develop a comprehensive analysis of financial crises, several streams of economic thinking can be combined: Knut Wicksell’s cumulative process, John Maynard Keynes’ analysis of uncertainty, expectations and herding, Hyman Minsky’s analysis of debt quotas and changing leverages, and Irving Fisher’s analysis of goods market deflation and his debt-deflation theory. Approaches in the tradition of behavioural finance can substantially add to the understanding of herding and various kinds of cumulative developments during periods of expansion and contraction.
Not all contraction processes lead to a financial crisis but there is always the danger that a contraction process gets out of control and develops into a deep financial crisis. The higher credit expansion and asset price inflation is, and the higher debt-quotas of economic units become during an expansion, the higher is the probability that a financial crisis will follow and the higher is the probability that the financial crisis will be severe. Financial crises can develop even without asset price bubbles. For example, when huge overcapacity is built-up during the expansion and/or households became highly indebted. It is obvious that the danger of a deep financial crisis is especially high when a strong economic expansion, which creates high overcapacity and increased debt quotas, is combined with a strong asset price bubble that is also, in the worst case, driven by credit expansion.

Financial crises typically develop after medium-term boom-bust cycles, which are usually not longer than one decade. However, there are also long-term developments over several decades possible which create fragile economic constellations. This occurs when the debt-income quotas of economic units increase over a longer period of time (see Hein/Dodig 2014). Given that debt-quotas are already high a medium-term boom-bust cycle with a systemic financial crisis must, then, be considered as especially deep. It also can bring the long-term expansion regime to an end. The subprime-crisis and the Great Recession is probably a historical example of such a “break” in a long-term development regime.

In the following section an own model of financial crises, based on the Keynesian paradigm including Minsky, is presented. The main drivers of systemic financial crises are, as mentioned, objective and subjective feedback mechanisms, which, during the creation of a boom or a bubble and during a contraction leading to a financial crisis, play an important role. A graphical exposition of a stylised expansion-contraction cycle with a systemic financial crisis is shown in Figure 2.

The beginning of an expansion is triggered by exogenous factors. This does not exclude the possibility that a period of positive economic development improves the state of confidence and expectations in such a way that an unsustainable development is triggered, which later leads to a systemic financial crisis. However, a period of low growth can reproduce itself permanently and can lead, over a long-time, to a very slow growing economy without an endogenous mechanism which could lead to an expansion.

The start of an expansion is characterised by a low interest rate, sufficient finance, and a high expected rate of return for investment in the enterprise sector. History decides which factors start an expansion and which will later lead to a bubble and a systemic financial crisis. From the financing side, a typical factor is a period of very low nominal interest rates (especially when interest rates were high before) or even negative real interest rates. Looking at the investment side a wide range of factors can trigger an expansion, for example, an innovation, a real shortage in production capacities or in housing, the end of a war, political developments, etc.

Objective feedback mechanisms are the goods market multiplier, which stimulates demand when investment increases. High investment via the goods market multiplier leads to high capacity utilisation and through the accelerator effect encourages further investment. High demand and especially demand inflation lead to high undistributed profits, which allow the self-financing of investment. This can further stimulate investment.
During an expansion banks reduce credit rationing and advance a high volume of credit. The uncertainty premium of banks is low. Private wealth owners reduce their liquidity hoarding and also start to expand credit to the enterprise sector and/or buy shares or other assets. During the expansion process asset prices start to rise. This triggers very strong feedback mechanisms in the financial sector. As the value of collateral increases, banks are willing to give more credit. Higher collateral also stimulates credit demand. Firms with increasing collateral feel it is safe to take on more credit. Additionally, households are able and willing to take more credit to increase their consumption, or to buy more assets, or to start speculating in asset markets.

During an expansion a whole set of subjective feedback mechanisms exist. Animal spirits or the state of confidence improve during the expansion. Optimism turns into over-optimism. Information cascades and stories which justify the development lead to herding, which further increases asset prices. Different groups of economic actors can have different expectations. Positive expectations of managers, expressed in a high expected rate of return for investment, lead to high investment. Such investment processes take time and lead to an expansion of real GDP and employment. Wealth owners including financial institutions may even have more positive expectations. Asset prices start to increase. As, for example, stock prices have no objective anchor, they can increase to very high levels driven by high current profits of firms and fantasies of
future profits. A clear sign of euphoria is increasing asset prices to levels not imaginable before. Pure speculation is triggered. Agents who do not have a long-term horizon enter asset markets. They buy assets even if they believe prices will fall in the long-run. They are convinced they are clever enough to sell before the mass of investors and speculators do so. Risk-loving investors speculating with credit also enter the market. Fraud and greed can become widespread. High asset prices further stimulate investment in the firm sector. High share prices allow companies to get cheap finance by issuing new shares. High real estate prices make it profitable for real estate developers to build houses and sell them in the market. Also, consumption demand is further stimulated by high asset prices.

During the expansion several instabilities build up. The most important point is that all sectors in the economy increase their gross debt. Within sectors indebtedness of certain units can become very high whereas other units do not accumulate more debt. High profits and the widow’s cruse may slow down increasing indebtedness of firms. Nevertheless, other firms may build up a high indebtedness or some private households become highly indebted. Others increase their net asset position and have no debt or debt is a small part of their assets. An overall extension of balance sheets takes place. Typically, asset price inflations are accompanied sooner or later by goods market inflations. The latter are driven by excess demand in the goods market and typically by a developing wage-price spiral. As long as the central bank does not increase interest rates, inflation can further stimulate the expansion.

What is important here is that the economy moves into a more and more fragile constellation. Indebtedness of all sectors in the economy increases and so do debtor- and creditor-risk. Asset prices increase to very high levels and become less and less sustainable. High investment in production capacities (including real estate markets) makes it less and less likely that all capacities in the future will be able to be utilised. It becomes clear that the expansion must eventually come to an end. Many factors can stop the expansion. For example, monetary policy may increase interest rates, companies in some economic sectors may realise less profits than expected, asset prices may stop increasing further and this may lead to problems of some speculators, or expectations change without any narrow economic explanation.

During a contraction period the problem of non-performing loans develops. When speculative activities in stock or real estate or other asset markets financed by credit do not materialise, or when investment projects of firms do not earn the cash flows expected, or private households cannot pay back their debt, either because asset prices do not increase further or income increases are below expectations, financial difficulties spread. Objective and subjective feedback mechanisms stimulate a cumulative contraction process.

An important objective feedback mechanism is the interaction between the fall in asset prices, financial problems of speculators and fire sales of economic units with liquidity and solvency problems. Fire sales can lead to an extremely quick and substantial collapse of asset prices and cause additional problems of speculative units. The more firms, financial institutions and households sell assets to be able to fulfil financial obligations, the faster asset prices will decline. Falling asset prices trigger further feedback mechanisms. They lead to a fall in the value of collateral and a destruction of equity. Banks will intensify credit rationing and will increase their uncertainty premium. Credit rationing may be enforced by a destruction of equity in
balance sheets of financial institutions and a decrease of collateral of debtors. In addition, the state of confidence erodes and the level of uncertainty increases. Banks follow a more and more cautious attitude to grant loans. A credit crunch can follow including even the breakdown of the short-term credit market between financial institutions.

Investment by firms collapses either because animal spirits erode or firms cannot get credit. Losses in the enterprise sector, caused by decreasing demand, further weaken investment. Now the goods market multiplier and falling consumption caused by negative income and wealth effects intensify the reduction in output and employment. Reduced aggregate demand leads to increasing financial problems in firms and to bankruptcies. Unemployment and shrinking household income lead to increasing non-performing loans for consumption and real estate credits.

In the centre of negative feedback factors is the permanent erosion of wealth, confidence, herding and finally panic. Financial institutions, during a crisis, will reduce credit supply radically whereas hoarding of liquidity by wealth owners add to the liquidity problem in the economy. In such a situation liquidity in the market dries out. Economic agents thirst for money like a poor soul for relief. Euphoria gives way to over-pessimism and panic. Without the central bank taking on the role as a lender of last resort, the financial system is in danger to collapse in a systemic crisis. Before Walter Bagehot (1873) wrote his Lombard Street and before central banks understood their function as lender of last resort this is what happened. Almost all financial crises in the first half of the 19th century led to the collapse of large parts of the financial system.

A lack of goods market demand and a shrinking of production and employment can lead to deflationary tendencies. As soon as unemployment becomes very high or labour market institutions are weak, the possibility that nominal wages start to fall increases. As soon as unit-labour costs start to decrease a cost driven deflationary process is triggered. A sharp deflation under the condition of high debt is the most brutal objective feedback mechanism, which tears an economy into a widespread systemic financial crisis. This is so because a deflation increases the debt burden of all debtors in the domestic currency. In addition to this, deflationary expectations lead to a collapse of investment by firms, which do not want to compete with other firms that buy productive capital later for a lower price. Households also reduce their consumption demand for durables if deflationary expectations spread. Production and employment reduce further and additional nominal wage cuts can follow as unemployment deepens. The central bank becomes not much more than an observer since monetary policy cannot cut nominal interest rates below zero and even if the central bank offers high liquidity, banks do not increase credit expansion when they follow credit rationing.

5.3. Extension of financial crisis models to the international sphere

The above discussed model of financial crises provides an important element to understand international crises. As in the domestic economy there are boom-bust cycles on an international level (see for example Williamson 2005).

\[23\] O God, you are my God; early will I seek you: my soul thirsts for you, my flesh longs for you in a dry and thirsty land, where no water is (Psalm 63:1)
The precondition of international boom-bust cycles is the existence of international capital flows. With strictly regulated international capital flows, international debt crises cannot develop. After international capital flows began to be deregulated, many countries were affected by boom-bust cycles, which led to severe financial crises. Examples are: the Latin American crises in the 1980s, the Mexican crisis in 1994, the Asian crisis in 1997, the crises in Argentina and Turkey in 2001, and the crises in Hungary and other Central and Easter European countries in 2009.

In short, boom-bust cycles on an international level are very similar to boom-bust cycles on a national level. They typically develop along the following sequence (Herr 2011a). International capital flows become deregulated and/or there is a change in expectations about the future economic development of a country. These changes play the same role as “displacements” which can trigger domestic bubbles. Also, stories and new era ideas play an important role in international boom-bust cycles. For example, in the 1970s, the Latin American countries were considered to develop very positively and to catch up quickly with the developed world. In the 1990s, Asian countries “emerged”, in the eyes of international investors, as regions with almost unlimited potentials.

When a country is considered to be a good investment location, domestic and foreign entrepreneurs, financial institutions and wealth owners typically share the same positive expectations. Capital starts to flow into the “emerged” country via different channels. Firstly, domestic financial institutions and bigger enterprises take foreign credit denominated in foreign currency. A domestic credit expansion is usually further stimulated or even driven by foreign credits. Portfolio investment in form of debt securities or shares and foreign direct investment further stimulates domestic development. The domestic expansion leads to all positive feedback mechanisms and phenomenon which were analysed in the sections above. In principle, therefore, the build-up of fragility during the expansion is identical with the one in a purely domestic model.

However, there is an important difference between a purely national and an international boom-bust cycle. In an international boom-bust cycle external fragilities join domestic fragilities. Net capital inflows lead to current account deficits. However, during a boom the lack of demand stemming from such deficits is compensated by dynamic domestic demand. But the key point is that foreign gross indebtedness increases. Most countries in the world suffer from the so called “original sin” (Eichengreen/Hausmann/Panizza 2003), indicating they are not able to take foreign credit in domestic currency. Their currencies are simply not trustworthy enough to be used as international standards for credits. Only countries which issue currencies at the top of the currency hierarchy like the United States or the Euro Area countries are able to accumulate foreign debt in domestic currency. Foreign debt denominated in foreign currency is extremely dangerous because a real depreciation of the domestic currency leads to an increase of the real debt burden in domestic currency. A real depreciation in countries with high debt in foreign currency has a comparable effect as a goods market deflation in domestic currency – in both cases the real debt burden overall increases and adds to the overindebtedness in a systemic way. Many developing countries suffer from high dollarization. When domestic loans are denominated in foreign currency, the same problems of currency mismatch, as mentioned, are created (Herr 2008). A boom phase connected or driven by the accumulation of foreign debt leads to three main fragility areas: an expansion of domestic credit and indebtedness in domestic currency, asset price
inflations, and the build-up of foreign debt and a currency mismatch in the enterprise, financial and sometimes even the household sector.\(^{24}\)

When a boom in combination with foreign indebtedness comes to an end, the bust phase usually becomes more pronounced and destructive than a boom that is only domestic. The explanation is that the bust is typically connected with a sudden stop of capital inflows and capital flight of foreign and domestic wealth owners. In addition, foreign debt has to be rescheduled. The country has a high need for foreign currency but simply has no access to it. When foreign reserves are depleted, the exchange rate collapses and in a situation of high foreign debt a twin crises becomes unavoidable (Kaminsky/ Reinhart 1999). And all this happens in addition to a collapse of domestic asset prices, non-performing loans in domestic currency and a shrinking domestic production and increasing unemployment. The problem is that the central bank and the government have no means to solve the foreign debt crises as they have no possibility to obtain foreign currency. The domestic central bank cannot take on the role as a lender of last resort because it can only issue domestic currency. Their only possibility is to approach, for example, the International Monetary Fund or other foreign institutions and comply with their dictated adjustment programs.

There is one further interesting point. A foreign financial crisis in one country can spread to other countries through different channels. When a financial crisis from one developing country spreads to another, it is usually not through the direct trade channel. And developing countries do usually not have excessive credit relations with each other. Even under these conditions financial crises can spread. The contagion channel is a change in conventional judgement of the risks investing in developing countries by the international investment community (for empirical examples see Kindleberger 1996).

After the deregulation of international financial markets, starting since the 1970s, several worldwide boom-bust cycles between the capitalist centre and the periphery have been observed. Capital flows from the centre to the periphery when the centre is booming and they flow back when the centre is in a crisis. This phenomenon creates crises in developing and developed countries at the same time. The first wave of such global capital flows started in the 1970s and came to an end in the early 1980s. The next wave began in the early 1990s and ended with the Asian crises in 1997. The last wave developed during the 2000s and came to its end with the subprime crises. Such waves destabilise the world economy, especially because a powerful international lender of last resort is lacking (Kindleberger 1996, Williamson 2005).

6. Policy Conclusions

From our analysis for economic policy the following conclusions can be drawn. Firstly, economic boom phases combined with credit expansions and higher indebtedness of economic units are the key explanation of a later systemic financial crisis. To prevent financial crises, such credit expansions have to be prevented. Microeconomic prudential regulations are important, but not sufficient to prevent an unsustainable

\(^{24}\) Public households also can take credit in foreign currency and can become over-indebted. We do not discuss the problems connected with high foreign public debt.
expansion. The individual banker, the individual firm or the individual buyer of real estate may all believe that they act in a responsible and rational way in spite of an unsustainable credit expansion. There must be a supervisory agency and supervisory mechanism on a macroeconomic level. Investment in industry, real estate, current account constellation and asset price developments have to be supervised carefully. Such a supervisory agency has the difficult task of analysing macroeconomic developments and distinguishing between expansion which create production capacities and jobs, and increase welfare and those expansions that can lead to costly financial crises.

Secondly, short-term speculation in asset markets, which are often linked with credit expansion and fraud, has no positive social function. Measures should be taken to reduce this type of speculation. Credit for speculative purposes should be strictly prevented. In case of overheating in the real estate sector or other sectors, the central bank should have instruments to reduce credit expansion in a selective way. Also, transaction costs in asset markets should be high. In addition, speculative gains should be taxed.

Real estate markets are very important for economic development and also for economic destabilisation as large credit volumes are involved in real estate activities. At the same time, the real estate market has a fundamental social function in society. Real estate markets should not be fully integrated in financial markets and should be subordinated under special regulations.

Goods market deflation creates one of the strongest destructive feedback mechanisms in financial crises. As soon as the nominal wage anchor breaks, deflation can develop in an uncontrollable way. To prevent deflation driven financial crisis labour market institutions must be strong enough to hinder nominal unit-labour costs from falling also when high unemployment prevails. In addition, demand management is needed to prevent a demand driven deflationary process.

To stop international boom-bust cycles, which are especially destructive, capital controls are needed. These can take different forms. One important area is financial market supervision to prevent any relevant currency mismatch. Economic units with no hard currency revenues should not be allowed to take foreign currency loans. Another area is to control capital inflows. Important for the stability of the world global financial system is a new global governance system, which includes the establishment of stable exchange rates, mechanisms to prevent high current account imbalances and the cooperation between key central banks like the US Federal Reserve and the European Central Bank (Herr 2011a).

To compress financial bubbles completely is difficult. Supervisors may hesitate to stop boom phases with innovations and the creation of production capacities and employment too early. Also, it is sometimes difficult to forecast bubbles and wrong judgements cannot be completely avoided. Instruments and procedures are needed to solve a financial crisis in a quick and efficient way without creating moral hazard.

---

25 See for example Eatwell (2004) for the positive functions a bubble can play.
7. Literature


