The limits of arbitrage and financial crises – implications for banking regulation and accounting standard setters

Markus Grottke[1]; Thomas Wenger[2]

Abstract

We investigate the conditions under which the assumption of arbitrage free markets allows for pricing assets and identifying arbitrage opportunities concerning ill-priced assets and the conditions under which such objectives cannot be rationally pursued. In two sophisticated examples we demonstrate why the latter situation could well be seen as a normal case. Implications from our setting for the banking regulation and accounting, especially fair value accounting, are discussed subsequently.

JEL Classification: M40, M41, G1

Key words: arbitrage, fair value accounting, banking regulation

1 Assistant Professor, University of Passau, Chair of Taxation, 94032 Passau, Germany.

2 Assistant Professor, University of Passau, Chair of Finance and financial control, 94032 Passau, Germany.

1. Introduction

There have been many explanations for the financial crisis from outside of finance. Less often, it has been tried to explain the financial crisis from inside, that is, by investigating the most common structures embedded in the models market participants regularly apply. This is what we will engage in.

Our analysis centers on the opaque character of the assumption of arbitrage free markets. It allows for the following contributions. First, we show that arbitrage free markets are only one extreme in a whole continuum of possible pricing processes. While the assumption of arbitrage free markets allows for establishing pricing processes that in turn allow for deriving prices from the market structure that can be applied to other assets and that allow for identifying ill-priced assets and applying transitional arbitrage strategies, nothing can be said about the impact of arbitrage strategies in case they interact. When arbitrage strategies are persistent, their presence could even cause financial crises and severely distort capital allocation. This insight points to the fact that the conditions of arbitrage free markets cannot be assumed or tacitly embedded in valuation procedures but have to be thoroughly checked. We demonstrate in two practically highly relevant examples of asset substitution that it is often impossible under the given conditions to fully guarantee those conditions. As a result, several consequences are drawn for the banking industry and for the viability of central objectives of "fair" value accounting.

We will proceed as follows. In the second section we provide a short literature review about the different explanations that have been provided for the recent financial crisis. In the third section, we develop step by step a simple example about what might be a common abstract structure that leads to modern financial crises. In the fourth section we demonstrate how this simple example evolves under more complex circumstances. We provide evidence why it cannot be tackled with the conventional remedies applied at the moment, be it regulatory devices, be it accounting rules. In the fifth and the sixth section, we raise the question which steps have to be taken to improve the current regulation of financial institutions and the accounting environment to reduce the likelihood of future financial crises. We conclude with a summary of our main results.

2. Literature Review

The prediction that the financial crisis provides us with research material for years (Roberts and Jones, 2009) has already proven to be true. There have been so many different explanations for the subprime crisis that the crisis itself has become multifaceted.

In the finance literature, the crisis has been explained as the result of securitization processes, exogenous credit supply shocks, or an interaction of different factors influencing the lending process.

The first strand of the finance literature explains the crisis by the new instruments of securitization. Gorton (2008) shows for example that the securitization chain was, in principle, highly sensitive to house prices. However, capital markets did not perceive the risks as complex securitization structures created asymmetric information. He explains the crisis as a panic that took place when the transparency introduced via the ABX indices suddenly allowed to aggregate the respective information about the securitized assets leading to subsequent negative reactions of market participants. Keys, Mukherjee, Seru & Vig (2010) add empirical evidence that securitization practices led to increased incentives to screen borrowers badly. However, Ghent (2011) provides a meticulous examination of the foreclosures of loans in the Era of the Great Depression. He demonstrates empirically that in an environment characterized by very similar circumstances to the subprime crisis (average nominal interest rates fell, average original maturity of mortgage loans was short, home prices rose and then experienced a sharp fall), however, without securitization, similar effects as in the subprime crisis were obtained.

A second strand of the finance literature explains the financial crisis on the basis of credit supply effects. Mian and Sufi (2009) test three different hypotheses that could potentially have caused the financial crisis: the increase of the income of subprime borrowers in the period before the crisis (income hypothesis), an external credit shock increasing the supply of credit (supply hypothesis) and finally, the expectation hypothesis that explains the crisis as a result of credit issued in expectation of increasing house prices. They can reject the income hypothesis because they provide empirical evidence that the relative income of subprime borrowers declined. Moreover, they show that the supply hypothesis is best in explaining the expansion of mortgage credits, while the expectation hypothesis cannot explain why credit supply did also increase in areas where house prices did not increase. They demonstrate further that there has been a sharp increase in the number of subprime mortgages on the one hand and a sharp decline in the prices of those mortgages on the other hand. Simultaneously, a decrease in non-mortgage debt is shown to have taken place. Similarly, Gârleanu and Pedersen (2011) demonstrate in a model with heterogeneous-risk-averse agents facing margin constraints how deviations from the Law of One Price could take place. Among other things, they show that reduced margins offered by the central bank could lead to higher asset prices which in turn could allow unviable loans to survive instead of defaulting.

A further strand of the finance literature focused on the conjoint effect of several factors. Demyanyk and van Hemert (2011) provide empirical evidence that the subprime loan quality deteriorated for six consecutive years. They provide evidence that the four variables that explain most of the delinquency are the FICO Score, the LTV Ratio, the mortgage rate and the house price appreciation. Khandani, Lo and Merton (2009) provide evidence that systemic risks in the financial market system are increased by the influence of three factors when they occur together, namely, rising home prices, declining interest rates and near-frictionless refinancing opportunities (or cash-out-opportunities). In extreme cases their simulation provides evidence that borrowers' cash-out-refinancing has the same effect as if all mortgages were originated at the peak of the housing market. However, they do not model the decisions of lenders which have – according to Mian and Sufi were a decisive variable in empirically explaining the subprime crisis.

Within the accounting literature the financial crisis has been also explained by accounting effects. One strand within the positivist mainstream literature explained the financial crisis by conjoint effects of management compensation schemes and accompanying accounting norms. Kothari and Lester (2012) e.g. explained that senior management had a strong incentive to "maximize the number of mortgages issued and sold to third parties."(p. 345). According to them accounting rules for securitization allowed for booking the profits early and overstate the assets subsequently, thus, positively influencing the benchmarks on which their own compensation was based (see for a similar explanation also Schildbach and Grottke, 2009; Laux and Leuz, 2009). Others pointed out that fair value accounting aggravated the crisis due to its pro-cyclical character which induced banks to increase their lending in the boom period before and to stop lending in the crisis afterwards (Plantin, Sapra and Shin, 2009). However, this explanation was contested from other authors from the mainstream literature. They pointed out that the financial crisis was at least in part the result of a lack of fair value accounting: The historical cost model provided only outdated information of deteriorations in the loan portfolios of failing banks (Linsmeier, 2011; Laux and Leuz, 2009). Therefore, fair values could deliver a proper remedy to this situation.

Especially the critical school of accounting provided some noteworthy explanations, dealing with the question why market participants did not see the crisis come. Gendron and Smith-Lacroix (2012) explained the crisis by the relatively weak level of paradigmatic diversity in finance which in turn leads to limitations of understanding the complex reality financial markets are confronted with. They point out that in finance practice and in the finance academy there are little indicators that finance is going to change its future course of action as there are no signs that an increase in paradigmatic diversity is fostered. Similarly, Roberts and Jones (2009), applying Latour's actor-network theory, explained the credit crisis as the result of the non-human agents, especially mathematical models and accounting for securitization, which enabled market participants to frame reality in a way that allowed for calculating and booking profits, while the true economic reality remained invisible in the models.

Taken together, the explanations of the financial crisis provided so far, prove to be heterogeneous and disintegrated. This makes it difficult to learn from them how to avoid future crises. Our question is therefore,whether a common ground of the crisis could be found and if so, how to make such a common ground sufficiently abstract to be applicable to a variety of potentially unsustainable structures within financial markets, while demonstrating that at the same time it can be translated into concrete structures to be observed in finance practice. Our work contributes towards that objective, as we take as a conceptual background mindset a combination of the following two frameworks, both of which were developed before the financial crisis and which deal with questions of the relation between information dynamics and stability properties of equilibration processes in considerable generality:

(1) the question of the reliability of model calibrations and its interplay with market efficiency or existence of equivalent martingale measures has been addressed by Rebonato (2004). He provided as a minimum recommendation to employ market implied valuations only for instruments which are actively traded in a market in which the assumption of fast corrections of transient ill-pricings can be plausibly made, i.e. when we may assume that arbitrage opportunities will dissipate quickly whenever they arise.

(2) One implication of the (vastly general) framework of H. Haken's synergetics (cf. Haken, 2005) is that a key condition responsible for the feasibility of computational treatments of information-driven equilibration processes consists in the validity of the principle of adiabatic elimination of fast variables. This principle is founded on an axiomatization of the conditions under which thermodynamic equilibration works well, i.e. leads to non-turbulent dynamics.

While a full application of the joint force of these two frameworks is beyond the scope of this paper, we develop the main ingredients of our analysis of the core mechanisms responsible for financial crises as translations of these two basic insights into the aspects of risk management at financial institutions, and their accounting and regulatory regimes.

.

3. The hallmark of modern financial crises

3.1 Arbitrage free markets

Regularly financial market actors engage in calibrating models that try to explain market prices of financial instruments by a minimum of price processes. The hallmark or cornerstone of this calibration process consists in the assumption that arbitrage is absent. To facilitate the understanding, we will provide a parsimonious example that is reduced to the absolute minimum necessary to understand what is happening. We assume that when the existence of our subsequently made argument is accepted for this most simple setting, it will be difficult to argue that the problem does not exist because finance practice is characterised by much more complex structures and models.

Let us assume a financial market that consists of five different financial assets which are quoted at a certain price in t_0 and will provide a cash flow structure dependent on different states in t_1. This market may involve three possible future states 1, 2 and 3.

[insert table 1 here]

Each state may have the same probability to occur. Market participants know these assets and their future cash flow structure. Obviously, the simple assumption that the market has to provide consistent prices (that markets are free of arbitrage) allows market participants to reduce the information contained in the market to the information of three single price processes of so-called Arrow-Debreu assets:

[insert table 2 here]

We can verify that this market is complete and arbitrage free. In other words, any other asset can be reproduced by these three price processes. As a consequence a unique price functional based on the three Arrow-Debreu price processes exists that fully characterises all financial instruments of this market:[3]

$$P = \pi(CF) = \sum_{i=1}^{3} p_i, t_0 \cdot CF_i$$

3.2 Introducing arbitrage

Now we assume that market participants are confronted with an additional financial instrument A 6:

[insert table 3 here]

Let us assume that a considerable number of market participants is well-informed, that is, they are sure about the process which is ill-priced. Then transitional arbitrage occurs, because informed market participants will engage in arbitrage, while the uninformed market

3 For a textbook treatment of the conceptual framework of arbitrage-free pricing we refer to Björk (2009).

participants will successively learn about their ill pricing and adapt (e.g. Grossman and Stiglitz, 1980; for conditions under which informationally efficient markets are obtained under such circumstances, see Hellwig, 1982). When e.g. A 6 is ill-priced, informed market participants will duplicate A 6 with the help of the three price processes and engage in arbitrage until any difference between the price of A 6 and the price of the duplication portfolio vanished. They will, for example offer A 6 for a price of more than 2,3 (which is the price of the duplication portfolio) and provide ample supply while reproducing and, therefore, buying a portfolio built from 3 P 1, 3 P 2 and 6 P 3.Uninformed market participants will step by step absorb the information contained in the market prices of previous periods and, therefore, also price the assets correctly at some stage. In other words arbitrage will be transitional and it will extend the information contained in the arbitrage free market also to A 6. One could, in line with Gorton (2008), e.g., interpret the securitization processes as a form in which banks arbitraged the ill-informed investors in secondary markets by duplicating their mortgage loan portfolio and selling the duplication on the secondary market. When the ABX index was introduced, however, the information asymmetry vanished on which the arbitrage was based.

However, normally, reality is not that clear. An entirely different situation occurs, when market participants do not know which of the four prices is obtained by correct pricing processes. What happens when perhaps one of the three price processes is mispriced instead of A 6? The problem is, then, that it is this price process which jeopardizes the assumption of arbitrage free markets instead of A 6 and arbitrage should concentrate on this price process. When market participants concentrate still on making arbitrage gains on A 6, they will divert capital flows from the correctly priced assets into the ill-priced asset et vice versa. This provides a possible analysis of why shifts from non-mortgage credit markets to mortgage credit markets occurred before the crisis as mentioned by Mian and Sufi (2009). With hindsight, they have been shifts towards ill-priced assets.

The situation is even worse when market participants do not know which one of the four price processes is "ill-priced". Under such circumstances, they have four potential opportunities how to arbitrage, because they can construct four different "arbitrage free markets" (P 1, P 2, P 3; P 1, P 2, P 3'; P 1', P 2, P 3; P 1, P 2', P 3) and four different opportunities to arbitrage (A 6, P 2, P 1, P 3). The point is that for each of these markets and arbitrage strategies, the market structure provides a different story, a different "unique" price functional.The respective differences are:

[insert table 4 here]

From this perspective, the problem is that market participants hold different expectations about the direction in which they should engage to arbitrage. As a result, dependent on their "arbitrage free" calibration process chosen and on the determination where arbitrage gains could be expected, different market participants will arbitrage different arbitrages away. This is another version of Roberts and Jones' (2009) story that market participants engaged in rationalising and framing their reality by the use of mathematical models, because when market participants *assume* that one price process is wrong, they can easily identify arbitrage opportunities that allow for booking profits.

Therefore, the direction into which the market moves then depends no longer on a "true" underlying economics or on the consistency of the market but is instead strongly influenced by the beliefs of the largest group of arbitrageurs. Are they, however, also the most sophisticated market participants? If not, there will be a potential that the "wrong" assets are demanded for arbitrage reasons leading to price movements which with hindsight are well interpreted as bubbles or negative discounted value projects (for conditions under which rational bubbles may form, see Kaizoji, Saichev and Sornette, 2011 as well as Malevergne and Sornette, 2001).

How can we find out whether such a situation occurs? Obviously, it would be necessary to whether one or all of the price processes are not reflecting all available information but instead are a deficient information dynamics. We could, for example, find out that the value process of A 6 is different from its price process, e.g. because state three can be divided into two different states with probability 1/6 where only in one of these states 6 are obtained while in the other state 6 have to be paid. Whether this finer value process is absorbed by the market then depends on the relationship between the information dynamic and the underlying value process of the respective assets. Of course, this is true for every asset in the market. The extend to which ill priced assets and bubbles occur, is, then, dependent on the fineness of the information dynamics of the majority in relation to the value processes of the assets. The most important point at this stage might be that valuation models that are based on the arbitrage free assumption might be "creating a mutually self-referencing hall of mirrors that [is] only shattered when real defaults and correlations" (Roberts and Jones, 2009, p. 864) take place – in other words, the arbitrage free market assumption acts like a drug which calms market participants down about their informational deficits.

We note that if we refrain entirely from assumptions about which assets are mispriced or which agents are better informed – as we certainly should if the situation continues to exist at the finest scale of resolution where information is available and reliable, then we arrive at the situation that protracted arbitrage induces shifts in the price system that defy any explanation whatsoever. Once we have used up all informational advantages and opportunities to pick finer price information by including more assets, but still have a choice of picking different maximally consistent sets of “basis” instruments, we must accept that such a market does not process information in a meaningful way at all during the period in which it has those properties – model results or accounting figures based on the arbitrage free market assumptions will, therefore, resemble a process properly dubbed “garbage in garbage out” during that phase, in the following sense: while there may exist a way or removing **all** arbitrage opportunities within a sufficiently short interval, and while we have reason to assume that the profit motive makes it likely that they will be used up, unless we check whether the “dissipation” has actually occurred across the whole multi-asset market at the finest scale of resolution, we may be mistaken when we conclude from the dissipation of **some** inconsistencies that **all** of them are gone. To the contrary, new ones may have appeared in the very process of making old ones disappear.

As a consequence, we can know that the assumption of absence of arbitrage will have harmless implications at larger scales if and only if those periods where it is violated are negligible in some sense, and the task would then be to specify in which sense – a task which arbitrage pricing theory has rarely faced.

The challenge resides in the fact that a simplistic way of specifying “negligibility” will not cut it, since we may face a situation in which the resolution of some “local” arbitrage opportunities introduces new “global” ones. Our elementary example above shows that this is abstractly conceivable and possible. The following examples show that the scenario is far from pathological, at least in the sense that the subprime crisis contained a large variety of realizations of both –they were not negligible “globally”, while appearing to be so “locally”.

We arrived at this characterization of the source of the problem of market turmoil by means of an analogy with thermodynamics, as is common in econophysics: the assumption of absence of arbitrage up to negligible corrections corresponds to the validity of the assumption that “fast” variables may be eliminated for processes sufficiently close to equilibrium. That such an assumption is not automatically satisfied and typically requires consideration of interactions with several insertions is then a guess that can be easily transferred across the analogy. We show by example that it leads to an economically meaningful perspective on the problem of characterizing the speed of decay of arbitrage. The “fast” variables would correspond to the price fluctuations arising from arbitrage activity while it is possible and profitable. In thermodynamics proper, the relevant condition would be the validity of Boltzmann’s so-called “Stosszahlansatz” (Boltzmann, 1896), which amounts to the assumption that the effects of interactions of two molecules in a gas will have decayed before new interactions occur.

4. Some more sophisticated application areas

4.1 Mortgages and gambling

In the following we assume a flat interest rate structure with the riskless rate r and look at a bank investing an amount E in equity that finances an agent who has taken up a mortgage valued at M_0 to purchase real estate which is subject to market fluctuations. The bank refinances this purchase by providing a loan amount of B, where B is viewed as a traded bond in the sequel, to be thought of as either a loan priced at fair value on the bank's balance sheet or a part of the underlying asset pool of a securitization tranche. We assume that the maturity T is fixed. The balance sheet thus turns up to be of the following structure:

[insert balance sheet 1 here]

Following the structural modeling approach of Merton (1974), the equity position is to be viewed as a call option on M with a strike that is the face value of the debt at maturity. We make the assumption that the loan contract allows for asset substitution, i.e. the bank may replace assets with alternative risk profiles at the same price, e.g. M_0 gets replaced , that is, substituted by the bank by a portfolio built from M' and G, of total value M' + G. M' is another mortgage with a reduced face value, for example, while G is an extra gamble, financed partially by B. The bank has to pay a small periodic fee for G which provides a very large outcome in extremely improbable situations. It may be seen literally as a lottery, and may be assumed as priced fairly. Even if it is priced fairly however, its benefits allocate to E while its loss potential hurts B, and a transfer of wealth occurs.

The new balance sheet looks like the following:

[insert balance sheet 2 here]

After an interval of time, we can expect that E rises while B will lose value, and we make that assumption, which is an assumption of partial effectivity of the "return to arbitrage freeness". Very obviously, we construct by fiat a situation in which assumptions of absence of arbitrage are inhomogeneous across periods and agents. That is precisely the point.

E can be expected to appreciate and B to drop in price because E has a positive vega, being a call option, and the gamble increases risk. Conversely, B is hurt because its probability of default is increased, since the periodic payments are no longer part of the assets when maturity is reached. For this to hold true, the gamble needs to be an actual gamble, meaning overall asset volatility must increase after the substitution. This is an assumption we make.

We think of an arbitrageur who acts before that interval of adjustment has passed. His idea is to short B or buy E long before the market prices adjust, or to do both simultaneously. I.e. he buys a certain amount of E (the long position in E) and promises to deliver B at a predetermined price (say the current one) in the future (say tomorrow), knowing that he will be able to buy B tomorrow at a much lower price than today, due to the drop in value of B which comes when the market perceives (assumed to take place overnight) the change in value due to the gamble, and thus making the difference his gain, since he then (tomorrow) sells at the (higher) price of today, since he sets up the contract with his buyer (who will lose money) today. To immunize this strategy against market risk arising from movements in M', hedging is necessary, which offsets the sensitivity for value changes of E and B caused by the movements of M'. Without such a hedge in place, the short position in B would lose value if M' goes up. Such a loss would be partially compensated by a gain in E, and in the event of a downward movement of M' it would be the converse, i.e. the short position in B would develop an extra pay-off and the long position in E would lose money. To cancel the risks precisely, one has to use a dynamically adjusted amount of M'. The original risky debt B corresponds to a short put on M' with strike equal to the face value F of the debt B at maturity T. The hedge portfolio is then determined by the put call parity:

$$B + M' = E + F \cdot e^{-rT}$$

The short position in B, hence, corresponds to the long put on M' with strike equal to F at maturity T. The required amount of M' to hedge the arbitrage portfolio is, hence, obtained from the put-call-parity by taking the sensitivities with respect to M'.

So, shorting B and hedging with M', for example, will produce a profit as the market adjusts B downwards. Similarly, buying E and hedging the M'-dependence away will produce a profit as markets include the increased volatility in their pricing due to G, which implies that E appreciates – the call option has a positive vega. But if we leave the strategy at that, there is some risk left, i.e. it is not yet an arbitrage strategy. In order to obtain both sources of profit in a riskless manner, one needs to also take the reverse position $G^{-1}ofG$ i.e. to buy the periodic stream long and let go of the pay-off of the gamble in the winning case. This allows for closing out positions after the gains from a rise of E and a fall of B are realized. When these close outs are not implemented, the arbitrageur retains positions that are subject to potential adverse value changes, i.e. we would no longer be considering an arbitrage strategy. As time passes, E and B may well change value again for reasons unrelated to M', since the odds of the gamble paying off change, as well as the number of periods in which the period payments can hurt the probability of default of B.

The whole strategy is an elaboration of the well-known problem of asset substitution to the detriment of creditors: creditors are hurt by an increase in riskiness of assets (e.g. Leuz, 1997). Typically, informed creditors would avoid that loss by imposing covenants before giving the loan, or by requiring E to be large enough at inception, so that the sensitivity of the bond B towards the gamble is small, while the bank holding E has an natural incentive to avoid the losses on E . The situation gets worse as E becomes smaller, but as long as unrestricted risk profiles for G exist and are available, and are not constrained by covenants, the effect always exists.

In fact the strategy can be interpreted in terms of the above mentioned hallmark of our general introductory example when projected on the time scale: The creditors' bond is a well-priced asset at first which becomes a mispriced asset, however, in that the asset substitution at the same price level takes place which in fact is the first arbitrage (which is well known in the literature since long, see e.g. Leuz, 1997). Since markets which are informed about this asset substitution only at a later stage or not at all, they will not adapt their prices of shares and bonds immediately. In this way, new arbitrage opportunities between the market values of bonds and equity and the market value of the gamble and the new mortgage are created by the first arbitrage strategy.

What we point out, thus, is that asset substitution – when the loan contract allows for it – provides an arbitrage gain, while at the same time opening up a capital structure arbitrage opportunity of a higher order of complexity at the very moment when the asset substitution takes place.

The partial trade of shorting the bond B and hedging via M' remains an arbitrage strategy which arises from the asset substitution put in place by the bank. Here, it is the first arbitrage strategy itself that induces that the information dynamics applied by other market participants becomes outdated. It merely does not suffice to remove all ill-pricings invoked by the asset substitution. Only when information would flow instantaneously no other arbitrage opportunities would open up when certain arbitrage opportunities are applied. We note that as long as creditors do not determine all actions in their contracts and as long as markets are informed via fair values or historical costs in a periodic and not instantaneous manner, such arbitrage opportunities will be available when asset substitution is available. While it may appear that the example is artificial since we assume at the outset that the creditor acts sub-optimally – by allowing for asset substitution – in reality the situation is typical, because even when covenants are specified and put in place, there is still a limit to the effective risk resolution of their detail. And beyond that limit, the situation will be of the same type as our example. So the example may be stylized, but it is not pathological.

The point, we want to stress, is that any credit could be subject to the first arbitrage strategy as long as asset substitution leading to more volatile risk profiles is possible and to the second arbitrage strategy as long as one can duplicate the equity or credit position faster in the market than the market can adapt to the new situation. It is important to stress, that such a structure has many application areas:

The widespread practice of banks (see Chen Chen, Liu & Ryan, 2008) to securitize risky asset pools while retaining only the riskiest first loss tranche or providing credit enhancements to special purpose entity structures (assuming part of the risk embedded in the vehicle) do in fact nothing qualitatively different from what has been demonstrated in the example.

Even the strategy of many house owners before the subprime crisis started, which was to tell banks that they would buy the house to live in while in truth they bought it to speculate with it (e.g. Ryan, 2008) can be interpreted as a private (and of course fraudulent) version of buying in a respective gamble. In fact the whole example of this section could be carried out on the balance sheet of the home owner himself.

4.2 An example of the secondary market

In the following we again assume a flat interest rate structure and look at a bank that owns an asset pool A that could, e.g., consist of loans and securities. The bank finances this pool by B (e.g. savings or deposits) and has an amount E in equity. We assume that the maturity T is fixed and again assume a flat interest structure with the riskless rate r. The balance sheet, thus, turns up to be of the following structure:

[insert balance sheet 3 here]

Again the equity position is not only an amount put into the bank by shareholders but it is at the same time a call option on A with a strike that is the face value of the debt. Again, we assume that the debt contract allows for an asset substitution. This allows for replacing A by A' and a credit default swap on a bond B' (cds(B')). A' is a smaller nominal amount of the same pool as A. cds(B') consists in writing an insurance against the default of another bond B'. We assume that for the prices at inception, we have $E \ll B', B \approx B'$. Again the cds provides a steady but low income stream and a low probability of a high loss potential which would, when occurring, wipe out E and hurt B severely, i.e. diminish its recovery if B' defaults. Since B' defaults only with very low probability – by assumption – the probability of default of B remains almost unaltered. The incentive for taking on this catastrophic risk comes from the appreciation of the value of E viewed as an option. In total, we obtain the following balance sheet:

[insert balance sheet 4 here]

Again, after a short interval Δt we can expect that E rises while B loses value. Therefore, the idea is again to short B or/and buy E long before the market prices adjust (and to "pick up" profits, closing positions afterwards). To immunize this situation against market risk arising from A', hedging is necessary, which offsets the sensitivity for value changes of E and B caused by the movements of A'. Otherwise, the short position in B loses value if A' is rising while not necessarily fully compensated by gains in E.

The original risky debt B corresponds to a short put on A' with strike equal to the face value F of the debt B at maturity T. The hedge portfolio is then determined by the put call parity:

$$B + A' = E + F \cdot e^{-rT}$$

The short position in B, hence, corresponds to the long put on A' with strike equal to F at maturity T. The required amount of A' to hedge the arbitrage portfolio is, hence, obtained from the put-call-parity by taking the sensitivities with respect to A'.

To arrive at all possible profits in a riskless manner, one has to include a further hedge, for the close-out after realization of profits. This hedge consists of buying insurance against default of B' at market prices, i.e. taking on cds (B'). The the pay-offs of the portfolio are:

[Insert table 5 here]

This example represents the effects of asset substitution on the recovery of bank debt. In the case of too big to fail banks it is the formalization of the socialization of severe losses, which is a known externality. It has led regulators to introduce the class of systemically important financial institutions, and demanding higher capital requirements from them, in an effort to internalize these externalities.

In this case, the assumption of informed creditors is even less plausible a priori than in the first example, since the very rationale for regulation is to create a representative creditor, mitigating the collective action problem of the depositors (Dewatripont and Tirole (1993)).

However, the point we want to stress, is that an implementation of the demonstrated arbitrage strategy would require the replication of the asset side, which is not at all feasible for systemically important banks. The very difficulty of replicating banks' asset pools – which is intrinsic, to be sure, since it is the very rationale for the risk transformation provided by banks – immediately involves that market discipline has only limited strength when it comes to safe-guarding against capital shortfalls or external effects. The problem is that market discipline requires strong assumptions on the ability of arbitrageurs to do their work. Those assumptions clearly become less feasible with hard to replicate asset pools – and are extremely implausible for systemically important banks, since then the structure of the liabilities will also become near-impossible to replicate, and of course in reality bank liabilities will be much more complex than in the stylized example here.

In a nutshell: assuming that banks know their own arbitrage strategies the argument strongly suggests why it is **rational** for interbank markets to dry up, and why even the increased capital requirements for systemically important banks require a considerable amount of fine-tuning to come close to being sufficient – simply because they would have to be at the level of the game of replicating the asset sides and capital structures of large banks dynamically.

5. Consequences for the regulation of financial markets

Two important messages arise from our investigations. The first one is that regulatory bodies should not base their investigations on the assumption of models that assume arbitrage free markets. Indeed, the models of finance might serve for all kinds of rationalisation; they will always allow for framing the reality of assets in a favorable way with respect to the model world designed.

What regulators have to do instead, is to question whether the models applied provide a realistic description of the current situation at financial markets. We believe that this will be so only under extremely rare circumstances. However, when models provide a distorted picture of reality, market prices based on model assumptions and model prices do not reflect the economic values of assets hold and liabilities owed by banks. Then regulators should thoroughly investigate the types of loans and their backgrounds. Obviously, model assumptions about which type of arbitrage free markets exist, then could be as diverse as human creativity and the different subsets of financial instruments allow. As a consequence, model scrutinization of regulators should be as diverse as well. In other words, regulators should make use of their competitive advantage in comparison to market participants which consists in receiving information about the applied models from *all* banks. More practically, it might be promising to concentrate on any exorbitant deviation, any shift in the assets invested in, or any inconsistency between model assumptions of different market participants, because such events could provide hints where a seemingly lucrative arbitrage strategy is followed which should be scrutinized for its viability. One way to find strategies of the kind mentioned in the second example might be to search for any increase in demand in assets that could be understood as gambles that could increase risk profiles via asset substitutions.

Furthermore, it makes sense to learn from those that have found heuristics how to find out the seeds of financial crises very early. Within a macro-economic context, Bezemer (2009) demonstrated that the application of equilibrium models reduced the opportunities to see the crisis coming. As an alternative he demonstrates that a number of those authors that engaged in early warnings before the financial crisis started have in common that they relied on an alternative accounting model approach that models the economy based on balance sheets. Such models allow for scrutinizing where the counter positions of assets and liabilities are to be found and how sustainable those relationships are. They made e.g. visible that the US economic growth before the crisis was based on the private debt growth, thus, providing evidence that an unsustainable state of affairs was reached.

Second our investigation relates to the question of capital requirements. The basic gap in capital requirements at the level of sophistication of Pillar 1 in Basel II was that they were calibrated purely with the probability of bank default in view, and not with the severity of bank default. The third example demonstrates, first, why this was an incomplete set of regulations. It also provides a partial rationale for that incompleteness: the asset substitution that hurts the recovery leads to an arbitrage opportunity – as does the asset substitution that hurts the probability of bank default. Obviously, also regulators will be burdened too much when they are supposed to scrutinize the impact of asset substitutions that lead to arbitrage opportunities concerning the asset pools of large banks. On the other hand, legal prescriptions have to be strict enough to be carried out afterwards. Therefore, we believe, that a mixture out of regulation and enforcement should take place. This regulation should start with asset substitution on behalf of banks and oblige banks to aggregate any asset substitution which leads to different risk profiles of the above mentioned kind. Furthermore, ideally it would set limits to asset substitutions, e.g. in that any substitution is not allowed for assuming additional losses that concern the uninformed deposit holders. The fact that this is not feasible in practice only shows that banks may thrive to a non-negligible extent on model risk, which they may effectively externalize by practicing asset substitution at a level of resolution much finer than regulatory standards and measurement errors. It is hence reasonable to view capital buffers as being partially, and even primarily, insurance against model risk. The implication is clear: setting capital requirements properly asks regulators to solve a fine-tuning problem. It is for this reason that they must be equipped with the ability to restrict the model approval process in banks.

Most importantly, we find that the following alternative holds: regulators must either be given access to the performance analysis of arbitrage desks at full detail, and must be able to carry out its analysis together with its implications for the capital adequacy of all other banks in the sector, or else arbitrage activity cannot be tolerated at all as a part of the business model of banks, since the very funding of it on the balance sheet re-introduces new potentially open and potentially more complex arbitrage opportunities in the market. These opportunities in turn undermine the assumption required for consistent model calibration across markets, which obviously must include interbank funding markets themselves.

6. Consequences for accounting

In fact fair value accounting, be it mark-to-market accounting, be it mark-to-model accounting can be interpreted as a form of information dynamics which could suffer from the above mentioned deficits. Our analysis raises doubts whether the three levels of fair value accounting, that is drawing on quoted prices in active markets or drawing on valuation techniques when such prices are not available does lead to more transparent financial reporting. As our first example shows, the resulting market price would need to be arbitrage free to provide meaningful information, otherwise it is simply the result of market participants which exchange assets on the basis of their model assumptions (Roberts & Jones, 2009). As our second and third example show, fair values would only be able to suggest the impact of asset substitution as far as outside market participants could fully understand the asset substitutions that take place in the banks' balance sheets, which is extremely unlikely. In any case fair value accounting does not enable them to understand that the asset side of the bank has substantially changed, when two assets with identical fair value but entirely different risk profile have been exchanged. In this sense fair value is not only the messenger who is shot for his true message (e.g. Laux and Leuz, 2009; Ryan, 2008) – it cannot report changes in risk profiles as would be necessary. Therefore, the advice to "simply" book the appropriate loan loss reserves (e.g. Kothari and Lester, 2012) might be not of such great value, because the question remains open based on which model this appropriate loss reserve should be calculated. There is also not so much hope that accounting rules provide a fine information dynamics, as repeatedly commentators pointed out that market participants themselves do change the information dynamics from the finer to the coarse, e.g. switching from Level one fair values to the Level three fair values (e.g. Kothari and Lester, 2012).

Our second example demonstrates, moreover, that the idea according to which investors are not naïve so that they will discover the deficiencies in the information dynamics (Laux & Leuz, 2009) might be premature. The point is that the complexity of the transaction and the time necessary can be increased to a degree which simply does not allow anymore for making sensible investments on the basis of fundamental information.

The question is, then, where the cure lies as far as accounting is concerned. Basically, it was suggested to increase disclosure (Ryan, 2008). When we accept that what is expected by market participants from the future, depends on the frameworks in which they conceptualize this very future and when we accept that this conceptualization and the subsequent actions based on it could be always incomplete (Roberts & Jones, 2009, p. 863), rendering the market prices meaningless, then the question is whether financial reporting is able to anticipate any possible way in which the frames currently chosen could be incomplete. We suspect that this task is simply too ambitious for accounting, because accounting itself needs to be put in chains by the concepts in which it is framed (Hopwood, 1990). Therefore, we wonder whether investors are better served by a very prudent version of a lower of cost or market accounting (for a suggestion in this direction, e.g. Benston and Wall, 2005) as we do not believe thatarbitrage strategies can vanish fast enoughvia informational remedies such as enhanced disclosure on the level of financial reporting. We believe indeed that the traditional mixture of historical cost accounting for any business that is not entirely finished yet together with a lower of cost or market-principle and provisions that absorb any potential gain made by e.g. redistributing risk from bondholders to shareholders is a clearly better solution than fair value accounting. When we understand fair values and the fair value gains and losses as a result of activities where self -interested agents engage in drawing and rationalising a picture that serves their interests (Roberts & Jones, 2009), then the most important thing is to secure that these rationalisations allow only for profits when they have been realized, that is, all sensitivity to material risks has dissipated, while they allow already for losses when those losses shine up at the horizon. Often this alternative has been forgotten and the traditional way of accounting is reduced to an understanding that is "ignoring market prices" (e.g. Laux & Leuz, 2009, p. 832).

7. Summary

In the preceding sections we presented what we believe to be a cornerstone of financial crises. First, we demonstrated how a rational model of arbitrage free markets can turn into a model which may divert scarce capital resources into worse investments. Then we provided evidence how asset substitution processes lead to arbitrage strategies. Finally, we have drawn several implications that arise for the banking regulation as well as for the accounting for financial instruments.

References

Benston, G.J. & Wall, L.D. (2005). How should banks account for loan losses. *Journal of Accounting and Public Policy*, 24 (2), 81-100.

Bezemer, D.J. (2010). Understanding financial crisis through accounting models. *Accounting, Organizations and Society*, 35(7), 676-688.

Björk, T. (2009). *Arbitrage Theory in Continuous Time*. 3rd edition. Oxford: Oxford University Press.

Boltzmann, L. (1896). *Vorlesungen über Gastheorie, Band 1: Theorie der Gase mit einatomigen Molekülen, deren Dimensionen gegen die mittlere Weglänge verschwinden*. Leipzig : Barth.

Chen, W., Liu, C.C. & Ryan, S.G. (2008). Characteristics of Securitizations that Determine Issuers' Retention of the Risks of the Securitized Assets. *Accounting Review*, 83(5), 1181-1215.

Demyanyk, Y. & van Hemert, O. (2011). Understanding the Subprime Mortgage Crisis. *Review of Financial Studies*, 24 (6), 1848-1880.

Dewatripont, M. & Tirole, J. (1993). *The prudential regulation of banks*. Lausanne: Editions Payot.

Gârleanu, N. & Pedersen, L.H. (2011). Margin-based Asset Pricing and Deviations from the Law of One Price. *Review of Financial Studies*, 24(6), 1980-2022.

Gendron, Y. & Smith-Lacroix, J.-H. (2012). The global financial crisis and the possibility of substantive change. Working Paper. Quebec 2012.

Ghent, A.C. (2011). Securitization and Mortgage Renegotiation: Evidence from the Great Depression. *Review of Financial Studies*, 24(6), 1814-1847.

Gorton, G. (2009). The Subprime Panic. *European Financial Management*, 15(1), 10-46.

Grossman, S. J. & Stiglitz, J.E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3), 393-408.

Haken, H. (2005). *Information and Self-Orgnaization – a macroscopic approach to complex systems*. 3rd edition. Berlin Heidelberg New York: Springer.

Hellwig, M.F. (1982). Rational Expectations Equilibrium with Conditioning on Past Prices: A Mean-Variance Example. *Journal of Economic Theory*, 26(2), 279-312.

Hopwood, A.G. (1990). Ambiguity, Knowledge and territorial claims: some observations on the doctrine of substance over form: A review essay. *British Accounting Review*, 22(1), 79-87.

Kaizoji, T., Saichev, A. & Sornette, D. (2011). Equilibrium model of rational noise traders: bifurcations and endogenous bubbles. arXiv:1109.4726v1

Kothari, S.P. & Lester, R. (2012). The Role of Accounting in the Financial Crisis: Lessons for the Future. *Accounting Horizons*, 26(2), 335-351.

Keys, J., Mukherjee, T., Seru, A.& Vig, V. (2010. Did securitization lead to lax screening? Evidence from subprime loans. *The Quarterly Journal of Economics*, 125 (1), 307-362.

Khandani, A. E., Lo, A.W., Merton, R.C. (2009). Systemic Risk and the Refinancing Ratchet Effect. Working Paper. Cambridge.

Laux, C. Leuz, C. (2009). The crisis of fair-value accounting: Making sense of the recent debate. *Accounting, Organizations and Society*, 34 (6-7), 826-834.

Leuz, C. (1997). *Rechnungslegung und Kreditfinanzierung*. Frankfurt: Peter Lang.

Linsmeier, T.J. (2011). Financial Reporting and Financial Crises: The Case for Measuring Financial Instruments at Fair Value in the Financial Statements. *Accounting Horizons,* 25(2), 409-417.

Malevergne, Y. & Sornette, D. (2001). Multi-dimensional rational bubbles and fat tails. *Quantitative Finance*, 1 (5), 553-541.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance,* 29(2), 449-470.

Mian, A. & Sufi, A. (2009). The consequences of mortgage credit expansion: evidence from the U.S. Mortgage default crisis. *The Quarterly Journal of Economics*, 124(4), 1449-1495.

Plantin, G., Sapra, H. & Shin, H.S. (2008). Marking-to-market: Panacea or Pandora's box? *Journal of Accounting Research*, 46(2), 435-460.

Rebonato, R. (2004). *Volatility and Correlation*, 2nd edition. Southern Gate, Chichester, West Sussex: John Wiley & Sons.

Roberts, J., Jones, M. (2009). Accounting for self interest in the credit crisis. *Accounting, Organizations and Society*, 34(6-7), 856-867.

Ryan, S. G. (2008). Accounting in and for the Subprime Crisis, *Accounting Review*, 83(6), 1605-1638.

Schildbach, T. & Grottke, M. (2009). Lehren aus der Subprime-Krise für Rechnungslegung und Corporate Governance. In Haeseler/Hörmann (Eds.), Wien, *Rechnungslegung und Unternehmensführung in turbulenten Zeiten* (pp. 107-126). LexisNexis.

Tables and balance sheets

Table 1:
Asset prices and corresponding Cash Flows

Assets	***Prices at t_0***	***State 1 in t_1***	***State 2 in t_1***	***State 3 in t_1***
A 1	*0,6*	*1*	*1*	*1*
A 2	*7*	*30*	*0*	*0*
A 3	*1,3*	*0*	*4*	*3*
A 4	*24*	*30*	*60*	*30*
A 5	*1,34*	*2,1*	*3*	*1,5*

State probability:
⅓.

Table 2:
Arrow Debreu price processes in an arbitrage free setting

Price process	Prices at t_0	State 1 in in t_1	State 2 in t_1	State 3 in t_1
P 1	*7/30*	*1*	*0*	*0*
P 2	*1/5*	*0*	*1*	*0*
P 3	*1/6*	*0*	*0*	*1*

State probability:
⅓.

Table 3:
Price and Cash Flows of Asset 6

Asset	Prices at t_0	State 1 in in t_1	State 2 in t_1	State 3 in t_1
P 3	*4 3/10*	*3*	*3*	*6*

State probability:
⅓.

Table 4:
Changing price processes under alternative arbitrage free settings

Price process	Prices at t_0	State 1 in in t_1	State 2 in t_1	State 3 in t_1
P' 1	*2/3*	*1*	*0*	*0*
P' 2	*19/30*	*0*	*1*	*0*
P' 3	*1 1/10*	*0*	*0*	*1*

State probability:
⅓.

Balance Sheet 1:
Point of departure: Balance sheet before the asset substitution

Balance Sheet 1

Assets	Liabilities
M_0	*E*
	B

Balance Sheet 2:
Balance sheet structure based on differently adapted economic values after substitution

Balance Sheet 2

Assets	Liabilities
M‘	*E*
	B
G	

Balance Sheet 3:
Point of departure: Balance sheet before the asset substitution

Balance Sheet 1

Assets	Liabilities
A	*E*
	B

Balance Sheet 4:
Balance sheet structure immediately after the asset substitution

Balance Sheet 2

Assets	Liabilities
A‘	*E*
	B
cds(B‘)	

Table 5:
Default structure of the underlying and pay of profile of the arbitrage strategy

Underlying	Pay off Profile
No default	$+\Delta E+\Delta B-cds(B')=cds(B')$
Default	$-E+B'+B-cds(B')$