DIPARTIMENTO DI SCIENZE ECONOMICHE E AZIENDALI
UNIVERSITÀ DEGLI STUDI DI PARMA
Via J. F. Kennedy, 6 – 43125 Parma – Italia

LIQUIDITY COGNITION AND LIMITS OF ARBITRAGE

ANDREA MANTOVI
andrea.mantovi@unipr.it

GIULIO TAGLIAVINI
giulio.tagliavini@unipr.it

WORKING PAPER EP01/2017

Parole chiave: Liquidity; Limits of arbitrage; Hierarchy of Money.

This draft: April 28, 2017

Liquidity cognition and limits of arbitrage

Andrea Mantovi, Giulio Tagliavini
Dipartimento di Scienze Economiche e Aziendali, Parma

E-mail: andrea.mantovi@unipr.it; giulio.tagliavini@unipr.it.

Abstract. *An inquiry about the limits of arbitrage connected with liquidity cognition is set forth. The hierarchy of money (Mehrling, 2012b; Pozsar, 2014) is shown to provide sharp insights on the emergence of endogenous constraints which may limit the strategic space of professional arbitrageurs in empirically relevant contexts. A case study (Mitchell and Pulvino, 2012) sets the stage for our theoretical proposal. Our approach is conjectured to represent a fruitful line of progress for behavioral finance.*

Keywords: Liquidity; Limits of arbitrage; Hierarchy of Money.
JEL Classification: E51; G01; G02.

This draft: April 28, 2017

1 Introduction

The great financial crisis of 2007-2009 has dramatically re-established the manifold problem of liquidity as a major challenge for academics, policymakers and bankers. The inquiry about the nature and causes of liquidity dynamics has regained momentum (BdF, 2008; Tirole, 2011), and seems to shape fundamental lines of progress for monetary theory and policy (see for instance Mehrling, 2012a; Stiglitz, 2015), and more generally for the understanding of the structural macro-instabilities at the roots both of the Great Recession (with its tragic impact on wealth levels and distributions) and the subsequent recovery path. Behavioral finance is meant to play a role in such respects (De Bondt et al., 2008; Shiller, 2014).

The 'evolution' of the liquidity problem has been pointed out among the reasons of its elusiveness. "Financial innovation, to an extent, may have let market participants believe that they could, on an enduring basis, escape from the monetary constraint (the need for genuine cash) and that they could make do with the liabilities issued by other institutions to meet their liquidity needs" (Bervas, in BdF, 2008: "Bervas" henceforth). Furthermore, as Tirole (2011) puts it, risk management defines the "flip side" of liquidity management, and cognition[1] about counterparty risk may be limited by the fact that such risk may depend on the solvency of systemic banks, and, in turn, on the stability of the entire financial system. It is the aim of this paper to deepen such insights on the cognitive problems which affect liquidity assessment and management, possibly resulting in limits of arbitrage connected with endogenous liquidity constraints.

As is well known, the cognitive problems which may plague a financial institution are rooted in first instance in the difficulty in monitoring and comprehending the state of the business (the collapse of Barings in 1995 is an instructive case study), and secondarily in the cognitive cycle about counterparty solvency. Problems of the first kind, at least in principle, admit a solution; for instance, "it is important to define unambiguous risk limits for traders, and then carefully monitor their activities to make sure that the limits are adhered to" (Hull, 2007, p. 49). On the other hand, for problems of the second kind *no* solution may exist, on account of the inherent complexity of liquidity dynamics, which, among other things, may bias market practices. As pointed out by Caruana and Kodres, "the expectations of market participants about liquidity and their ability to monitor it also have an impact on liquidity itself" (BdF, 2008, p. 65). True, the interactions between expectations and cognition about financial phenomena have long been acknowledged (see for instance Soros, 2013, and references therein).

By now, behavioral finance represents quite an established theoretical approach to the inefficiencies of financial markets (Shleifer, 2000; Shiller, 2014) and the patterns of behavior which seem to challenge standard rationality or efficiency benchmarks, in particular the efficient market hypothesis (EMH). In such contexts, liquidity problems do not seem as yet to represent a major line of inquiry; it is our aim to try and shape a behavioral perspective on liquidity cognition in connection with the limits

[1] For a systematic discussion of cognitive economics see Walliser (2008).

of arbitrage (Shleifer and Vishny, 1997; "SV" henceforth), which represent one of the pillars of behavioral finance (Barberis and Thaler, 2003). The impact of liquidity conditions on arbitrage performance has been inquired in recent years on both normative and empirical grounds (see for instance Mitchell et al., 2007; Nashikkar et al., 2011; Teo, 2011; Bhanot and Guo, 2012; Mitchell and Pulvino, 2012). In order to tailor a behavioral perspective on the problem, we shall argue about the hierarchy of money (Mehrling, 2012b; Pozsar, 2014) as a cogent theoretical perspective on the 'physiology' of liquidity provision. The hierarchy of money can be traced to the vision of money as balance sheet record of debt relations discussed by Keynes in the *Treatise* (see Bell, 2001, and references therein).

Kahneman (2003) points out the tension between, on the one hand, the elegant and consistent, albeit "psychologically unrealistic", rational model of economic choice, and, on the other hand, the somewhat more realistic list of errors and biases which characterize real human beings. Such a tension has been seen destructively by those scholars who think psychology does not provide a coherent alternative to the rational agent model. Still, one may consider such approaches not as substitutes, but rather as *complements*. In fact, according to the Author, the aforementioned tension should be seen constructively, i.e., as a guide to the construction of maps of bounded rationality meant to frame the kaleidoscope of real behaviour. Along similar lines, it is our aim to argue that the cognitive problem of liquidity may be approached constructively, namely, with the aim of building behavioral insights on liquidity[2] cognition which may enlighten the sources of instability of liquidity dynamics, with the dramatic consequences we know too well.

Ever since the "theory" of the Great Depression set forth by Irving Fisher in 1933, a number of sources of financial instability have been discussed (a recent addition to the list, fair value accounting). Among the few stabilizing forces, arbitrage is meant to play a major role; unfortunately, substantial limits to arbitrage activity have been displayed amid the crisis. Mitchell and Pulvino (2012) address such limits, and set the stage for our considerations: in the phenomenology discussed by the Authors, "from an arbitrageur's perspective, seemingly long-term capital became truly short-term overnight" (ivi, p. 490). To a large extent, the present contribution is an interpretation of the "seemingly" adverb in the quote.

The plan of the rest of the paper is as follows. In section 2 we envision the relevance of endogenous constraints. In section 3 we introduce the limits of arbitrage. In section 4 we discuss the hierarchy of money, which we employ in section 5 to focus some limits of liquidity cognition and arbitrage activity. Conclusions follow.

[2] It seems interesting to notice the huge amounts of cash held by nonfinancial public corporations in the US in recent years (Dittmar, 2008).

2 Endogenous constraints

Constraints represent basic elements of economic problems. It may be even argued that an unconstrained problem should not be considered an economic problem, at least if one defines economic problems in terms of allocations of *ex-ante* scarce resources. The role of constraints in financial problems is well known to be more nuanced. In first instance, solvability of a financial institution does not constrain *per se* the expansion of its balance sheet. Furthermore, prices and allocations are far more unstable, and the intertemporal nature of financial goals (say, maturity and risk transformation) makes it difficult to pin down ex-ante explicit forms for the constraints at play. Arguably, the more significant exogenous constraints faced by financial markets participants are given by the physical microstructures of markets themselves, i.e. the actual networks through which flows (of incentives, claims and risks) are channelled, network effects emerge (Brunnermeier, 2009), and policies get transmitted.

The analysis of *endogenous* constraints seems to represent a significant driver of current research efforts; for instance, endogenous constraints connected with the dynamics of collateral have been recently addressed by Araújo et al. (2015), and in fact the instability of collateral value represents a pivotal aspect of the phenomenology we shall be interested in. Specifically, our focus shall be the endogenous nature of (typically unanticipated) liquidity constraints, which contribute significantly to the complexity of financial stability problems. The potential emergence of liquidity constraints in investment decisions is hardly clear ex-ante, particularly so in expansionary phases, and even more so in connection with the subtleties of the practice of financial engineering, which, according to Bervas, "gives investors a misleading sense of liquidity" (ivi, p. 126). To a large extent, liquidity constraints manifest themselves ex-post, as emergent phenomena. Goodhart points out that "there is a tradeoff between stock liquidity and maturity transformation" (BdF, 2008, p. 43), but also that such a tradeoff can hardly be represented by quantitative scales, and that it is difficult to compare the maturity transformation positions of two banks.

In such contexts, talk about the "rationality" of a financial strategy may be slippery. We may well call "rational" an expected utility model of investment or balance sheet management. Still, the somewhat elusive nature of the liquidity constraints to which the agents may be subject makes it somewhat ambiguous to call "rational" the application of such models to real decisions impinging on the liquidity of an entity. One may well sharpen the picture, and explicitly specify liquidity constraints to which expected utility maximizers are subject (as for instance in Brunnermeier and Pedersen, 2009; "BP" henceforth), and still observe destabilizing liquidity spirals emerging from perfectly rational behavior. Stock liquidity seems to represent emergent phenomena which can hardly be reduced ex-ante to the properties of entities.[3] Goodhart makes this point sharply: "There is no firm barrier on one side of

[3] Compare the simpler setting discussed by Holmström and Tirole (2011), in which liquidity problems are addressed at the entity level.

which all assets should count 100% for such stock liquidity and on the other side 0%" (BdF, 2008, p. 43).

Such a conceptual problem, which one may hope to confine to academia, gains dramatic empirical relevance with respect to the recent practice of "pushing back the liquidity frontier" (Bervas). Over the last decades, increasing confidence in the availability of market liquidity has resulted in a continuous change in the compositions of balance sheets, with a declining role in cash holding and liquid assets; a "social problem" of liquidity assessment (ivi) has been pointed out. The case study discussed by Mitchell and Pulvino (2012) provides explicit insights on the emergence of liquidity constraints, which we discuss in section 5.

Perhaps one should simply accept the possibility of unexpected liquidity dry-ups as unpredictable events against which only central banks, ex-post, can backstop (the moral hazard problem connected with such vision has been long appreciated). Still, our theoretical proposal is to envision relevant aspects of such a dilemma in terms of emergent scarcity of liquidity with respect to the "hierarchy of money" (see section 4), which seems to rationalize the apparent conundrum that, in the words of Mehrling (2012b), "liquidity is at the same time both scarce and elastic." As a first step in such direction, let us confront some cognitive limits which seem to have characterized agents' behavior in recent years.

2.1 Quantitative fallacies

Statistical fallacies are well known to be incumbent on empirical analysis. Once a firm understanding of a phenomenon is not quite assessed, statistics may be employed as a means to deepen such understanding; in such circumstances, the interpretation of data may be biased by the dataset itself. Biased interpretations may be conceived of track records, which may not represent sufficient statistics if the system has not been subject to the complete set of possible stresses (the stress tests to which banks are periodically subject are in fact meant to define one such sufficient statistics). For instance, Bervas points out that "In the long chain of securitization, some presumed risk absorbers proved in fact to be a source of distress contagion between markets." Such presumptions were motivated by the observable phenomenology; could one argue differently on account of unobserved phenomenology? No satisfactory answer seems to have been fixed as yet. Still, hopefully, the dramatic phenomenology of the last decade[4] may provide some good points with respect to which to deepen our interpretation of the criticalities of modern financial markets.

We are interested in particular in the interpretation of the *quantitative* variables which (seem to) characterize liquidity conditions at a definite date. The Economist (2008) strikes the point sharply. "In January 2007 the world looked almost riskless […] we could only see more liquidity coming into the market – not going out of it. Institutional investors, hedge funds, private-equity firms and sovereign

[4] According to John Authers, "We are all market historians now" (Financial Times, 2010, February 5).

wealth funds were all looking to invest in assets. This was why credit spreads were narrowing." Such misguided interpretations of the global outlook have been the subject of (ex-post) bitter criticism. In response to such criticism, the image of the "black swan" has been repeatedly employed as a justification of the inherent difficulty of assessing the probability of an extreme event like the 2007 credit crunch.

We seem to face the particularly misleading statistical fallacy according to which rare events can be considered as coming out of the blue – not from well defined transmission channels of distress – and then simply *disregarded*. Such approaches have been criticized in depth in recent years; still, a somewhat specific aspect seems to deserve further attention. The quote from The Economist (2008) focuses what we may call a *quantitative fallacy*, i.e. a misguided interpretation of the stabilizing role of the "quantity" of liquidity, which seems to entail *cognitive limits* about market dynamics and potentially emerging constraints.

As quoted in section 1, Bervas witnesses the widespread belief that liquidity was not a problem in early 2007. We can interpret this belief as limited cognition about the *micro* dynamics of market and funding liquidity, in particular, the role of dealers in demanding funding liquidity and supplying market liquidity. Such mechanics does not seem to imply that quantitative variables like the size and rate of growth of demand of a class of assets (say, "riskless" tranches of CDOs) may gauge the stability of the system. Furthermore, the potential emergence of liquidity constraints, sparking self-reinforcing market and funding liquidity spirals, can be modeled as in BP, so as to fix sharp insights on the problem of liquidity as inherently related to "non-quantitative" variables like margins and shadow cost of capital. Given that such elements were available before the 2007 crunch, an inquiry about the spreading of such fallacious beliefs seems to represent an insightful behavioral challenge, which we perceive as somehow connected with the arbitrage crashes we shall confront in section 5.

3 Limits of arbitrage

Recall, *ideal* arbitrage can be defined as "the simultaneous purchase and sale of the same, or essentially similar, security in two different markets for advantageously different prices"[5] and considered neither to entail risk nor to require capital. Such an ideal notion is one of the pillars of the EMH; at a somewhat more theoretical level, the "no arbitrage" condition is a foundational pillar of asset pricing methods (see for instance Cochrane, 2005). With respect to such ideal notions, limits of arbitrage and other market imperfections have been referred to as "anomalies." Still, behavioral finance has succeeded in displaying the 'physiological' character of such anomalies. Game theorists have long established that Pareto inefficient outcomes can occur as perfectly rational Nash equilibria; correspondingly, behavioral finance has succeeded in showing that inefficient allocations, limited

[5] Sharpe and Alexander, *Investments, 4th edition* (Prentice Hall, 1990).

arbitrage and instability may at times be perfectly rational market outcomes, as a result of best responses of agents to the emergence of (endogenous) constraints.

SV argue about the sources of risk which impinge on empirical arbitrage, in first instance, the fact that price differentials may not shrink in the short run, and eventually enlarge. That being the case, losses may at times be unavoidable if the arbitrageur is not concerned with holding positions indefinitely. Short run arbitrage gains on average. It follows that short run arbitrage requires capital. More specifically, SV deepen the agency problem faced by professional arbitrageurs who manage other people's money (like pension funds and hedge funds) and whose compensation is contingent on performance. The Authors set forth a model for the incentives driving investors, and establish that performance based arbitrage (PBA) can generate a critical phenomenon, namely, the fact that "when arbitrage requires debt or equity capital, arbitrageurs can become most constrained precisely when they have the best opportunities" (Vishny, 2000, p. 90). Such a theoretical argument reflects relevant empirical phenomenology; in fact, both the use of redemption gates by a number of hedge funds managers and the grant of favorable redemption terms by others (Teo, 2011) witness that redemption conditions represent a significant strategic variable for applied PBA.

This kind of phenomena represent our focus, which we shall deepen in connection with a seemingly relevant (and perhaps underestimated) aspect of the empirical analysis of arbitrage. On top of the limits of arbitrage discussed by SV lies the fact that the stylized notion of "essentially similar securities" provides a somewhat loose guidance to the empirical inquiry about the business of professional arbitrageurs (for which hedge funds represent examples *par excellence*). One faces evidences that securities are fundamentally *unique*, and their 'degree of similarity' is contingent on market conditions, expectations about underlying processes, and more importantly, the *specific* mechanisms meant to exploit arbitrage opportunities. "Low-hanging fruits" are typically scarcely interesting to professional arbitrageurs, who rather focus on sophisticated strategies concerning management of combinations of securities. For instance, convertible debenture arbitrage and merger arbitrage do not represent truly arbitrage strategies (Mitchell and Pulvino, 2012). A recent focus of empirical analysis is the riskiness of arbitraging ETF prices and their NAVs (Xu and Yin, 2017).

In such contexts, cognitive limits of arbitrage may emerge once such strategies expose the arbitrageur to unanticipated market instabilities (and then, typically, liquidity stress). For instance, the stabilizing role ascribed to MMFs before the crisis may be taken as an example of cognitive limits which have manifested themselves at the crisis unfolded. In such respects, the case study discussed by Mitchell and Pulvino (2012) about crashes of hedge fund activities shall define the terrain upon which to build our arguments.

The collapse of LTCM in 1998 may provide further empirical backing for the insights we are pursuing. As conjectured by M. Scholes: "Maybe the error of Long-Term was […] that of not realizing that the world is becoming more and more global over time" (as reported by MacKenzie, 2006, p. 242). Tirole (2011) seems to align with such a vision upon acknowledging that the solvability of the entire

financial system may at times be the dramatic horizon facing market participants. In essence, the more competitive and interconnected the financial system becomes, the less risk can be manufactured (Acharya et al., 2009), the more elusive liquidity dynamics becomes, with substantial consequences of arbitrage activity as a stabilizing factor. This is the vision which underlies our theoretical perspective, which we introduce in the following section.

4 The inherent hierarchy of money

In the last decades financial innovation has been evolving the interactions between financial and monetary phenomena; in particular, the increasing role of market-based credit has contributed further elements of complexity to liquidity dynamics. The relevance seems to follow of deepening the connection of behavioral finance with the 'physiology' of liquidity phenomenology, which plays a pivotal role in the stability of markets and wealth distributions. In our view, the inherent hierarchy of money (Mehrling, 2012b; Pozsar, 2014) represents an insightful perspective in such respects, as we try and argue as follows.

Standard monetary theory and central banks fix a number of *monetary aggregates* as classes of instruments which embody varying degrees of liquidity and moneyness (the two attributes can be taken to be monotonically related for our purposes). Currency displays the maximal degree of liquidity and moneyness, in that it *cannot be refused* as a means of settlement, and fixes the benchmark value with respect to which to define par. Instruments belonging to larger aggregates display lower liquidity or moneyness; for instance, bank money may not be immediately redeemable at a bank under serious stress, and MMF shares may not be useful for settlement purposes.

Monetary aggregates represent a natural and quite established hierarchy of liquidity and moneyness; still, a more profound vision on the inherently hierarchical nature of moneys is given by Mehrling (2012b) in terms of the *dichotomy* between the principle of elasticity and the principle of discipline. On the one hand, the principle of elasticity provides the rationale for credit expansion, once the owner of a certain quality of money assesses the profitability of some lending activity. On the other hand, the principle of discipline provides the rationale for credit control, i.e., not only the creditworthiness of borrowers should gauge credit expansion as common sense discipline, but the expansion of credit – more elasticity – *today* inherently implies a contraction of credit – more discipline – *tomorrow*, when promises to repay come due. The two principles can be considered the opposite faces of the same coin, sort of yin-yang pair; from an economic standpoint, there is tradeoff embedded in the dichotomy.

In fact, Mehrling (2012b) applies the elasticity-discipline dichotomy to unfold an inherent hierarchy of moneys ("*H*" henceforth) in which the best money of the system (central bank money for a closed system) stands at the top, and supports the creation of moneys at lower layers of the hierarchy as credit relations, in which "what counts as money and what counts as credit depends on where in the hierarchy you are standing" (ivi). For instance, a bank deposit counts as money for the depositor and

as credit for the bank. One thereby conceives of a pyramid of money layers in which a 'small' (scarce) amount of the best money of the system sparks a hierarchy of layers of 'inferior' monies: definite institutions (the central bank, commercial banks, etc.) straddle the layers of the hierarchy by extending credit to some borrowers, and thereby creating money at the adjacent lower level. Much like for monetary aggregates, liquidity and moneyness diminish with the distance from the best money of the system; true, the two-sided nature of money as credit provides a deeper account of the physiology of monetary phenomena, which seems to represent the crucial point for the interpretation of the great financial crisis of 2007-2009, as well as for the design and regulation (under way) of market based credit.

Noticeably, *H* seems to provide quite a natural generalization of the canonical view on the physiology of credit, as represented by the following statement by Shubik (2010), which we elect to denote as a proposition for ease of subsequent reference:

Proposition 1

"banks can create a near money that serves on par with government money or gold and gives them strategic power[6] and risk evaluation responsibility in a flexible credit system" (ivi, p. 237)

The "strategic power" to which Shubik (2010) refers can be interpreted as the option of an institution at a level of the hierarchy to strike a contingent balance (tradeoff) between elasticity and discipline in the provision of credit (creation of liquidity). It seems appropriate to state that *H* represents an increasing strategic power of the upper levels of *H*, which rests on "the type and quantity of liquidity reserves, and the type and mix of liquidity and credit backing them" (Pozsar, 2014, p. 10). It is our aim to exploit such a vision on the physiology of liquidity dynamics in order to deepen the arbitrage crashes which characterized the financial downturn, and more generally to broaden the reach of limits of arbitrage analysis.

The point is that *not* all credit relations can be represented along *H*. For instance, a friend of mine may lend me a thousand euros without worrying much about repayment horizons, or default on my (soft) obligation. In this case, a *personal* relation is involved in the imbalance between elasticity and discipline: friendship provides the rationale for generous elasticity; in addition, the fact that a thousand euros may not represent a significant loss for the lender, further weakens the bite of the discipline element. True, also *impersonal* (market) lending may not fit *H*, in particular once lending is not supported by a superior money (think of bank reserves as the paradigm of superior money supporting the creation of inferior money as credit). The evolving role of collateral (Singh, 2013) in backing the provision of liquidity represents a major source of credit expansion (and instability) not

[6] Shubik (1999) advocates a game-theoretic approach to the dynamics of monetary phenomena. Interestingly, such a methodological approach is consistent with the one employed by Walliser (2008) for describing the coordination of players through beliefs.

necessarily aligned with *H*. The complexity of the current financial panorama is still under inquiry (see for instance Pozsar, 2014; Grillet-Aubert et al., 2016); still, we seem to envision quite a robust conceptual standpoint in stating that *H* provides a cogent vision on the physiology and *stability* of liquidity dynamics, and that the analysis of the limits of arbitrage can be aligned with such a vision in order to sharpen our map of the fragilities of the system. In essence, in our view, the 'degree of alignment' with *H* provides a crucial indicator of the stability of such relation, and, in turn, of its consequences on the system, in particular if arbitrageurs are involved.

The relevance of such matters extends beyond academia. Both the European Systemic Risk Board (see for instance Grillet-Aubert et al., 2016) and the Financial Stability Oversight Council (2016) highlight the pivotal role of the interactions between entities (balance sheets) in gauging systemic stability, which is, by definition, a dynamical issue. In such respect, Pozsar (2014) argues about the relevance of mapping such interactions and fix a significant picture of market-based credit activity, as well as of defining a new set of monetary aggregates, whose focus should not be price stability, but rather, financial stability:[7] in the words of the Author, "for institutional cash pools, money begins where M2 ends". In particular, for institutional cash pools, money has been (and still is) represented to a large extent by MMF shares. MMFs "have been thought of as immune to runs and an important contributor to financial stability. Had not the subprime crisis occurred, this would perhaps still have been the prevailing belief" (Bengtsson, 2013, p. 580). In fact, Pozsar (2014) is quite explicit in ranking MMFs below commercial banks along *H*, and arguing that their services can hardly gauge the stability of liquidity provision.

Furthermore, we can deepen our understanding of the quantitative fallacy (subsection 2.1) in terms of *H*. The increasing liquidity quoted from The Economist (2008) can be interpreted in terms of the elasticity of the system, once increasing market liquidity calls for an increase in funding liquidity meant to lubricate the efficiency of the system in generating and allocating the assets demanded. True, the design of *H* enables us to appreciate in which sense *more elasticity today begets more discipline tomorrow*, in particular if asset prices happen to fall and spark a liquidity spiral. This is a sketchy albeit useful insight on the quantitative fallacy, and perhaps on the debate about the effectiveness of monetary policies meant to provide "quantitative easing."

Finally, we are in a position to employ *H* in order to enlighten some cognitive limits which may affect liquidity assessment and management, and in particular limit arbitrage activity.

[7] In such respect, we do align with Stiglitz (2012) in deeming financial stability as far more important than price stability.

5 Liquidity cognition and professional arbitrage

As market historians, we know that liquidity problems may emerge somewhat abruptly, in spite of seemingly stationary conditions, and result in liquidity spirals possibly interacting with deflationary spirals (for which Brunnermeier and Sannikov, 2016, provide a theoretical perspective). Liquidity is essentially a dynamical problem, and the crisis has been interpreted in terms of Minsky's cycles (see for instance Bengtsson, 2013; The Economist, 2016) which, arguably, reinforce the procyclical character of PBA activity as depicted by SV. Along such a line of thought, let us employ the phenomenology discussed by Mitchell and Pulvino (2012) as a playground for our theoretical proposal.

The Authors provide an empirical analysis of arbitrage crashes concerning prime brokerage debt financing of hedge funds. Such a mechanism has played a role in triggering the onset of the 2007 crunch: once changing market conditions led a number of banks to reduce their credit exposure to hedge funds, demand for mortgage backed securities began to decline; as a consequence, a number of investment funds and structured vehicles, devised to invest in credit instruments with short-term borrowing, came under severe pressure (Fisher, in BdF, p. 32). Such prime brokerage financing seems to provide a sharp empirical setting for our considerations.

According to the SV stylized approach, the limits to which PBA is subject stem from the responsiveness of fund availability to past performance: it is the agency relation between the stylized investor and the stylized professional arbitrageur which generates the emergence of endogenous constraints on the strategic space available to the latter. With respect to such a theoretical problem, Mitchell and Pulvino (2012) enlighten the nuances which enrich the agency problem once considering its empirical manifestations, in which investors' decisions may not represent the pivotal elements of the game. In fact, the differentiation of capital which characterizes a concrete arbitrage activity entails specific incentives driving the diverse suppliers of capital, and then, possibly, the sudden emergence of constraints on their strategic space. In turn, such constraints reverberate on the professional arbitrageur.

The business model of Lehman Brothers International Europe (LBIE) is considered, as a prime broker providing a bundle of services to hedge funds, among which loans at favourable rates generated by a mechanism (ivi, Fig. 1) which entails a rehypothecation of assets. A hierarchy is there represented as a lending chain. The first link goes from a rehypothecation lender to the prime broker, the second link goes from the prime broker to the hedge fund: the prime broker employs part of the collateral pledged by the hedge fund to collateralize the loans obtained by rehypothecation lenders. Such loans are a further source of capital which can be lent to the hedge fund.[8] Thus, a two-level chain ("*h*" henceforth) generates the leverage of the hedge fund. The efficiency of *h* was witnessed by

[8] Iterated collateralization has by now become a significant phenomenon: according to Singh (2013), the size of collateral reuse has become comparable with monetary aggregated like M2.

favourable rates (slightly above the risk-free rate); still, its stability has been severely challenged by the events of September 2008.

The filing of Lehman Brothers in September 2008 had an impact on the sustainability of *h*. Rehypothecation lenders were forced to sell the securities pledged as collateral; LBIE clients became unsecured creditors of LBIE. A delinking of portfolio positions occurred in the levels of *h*: contrary to the hedged positions of the hedge fund, the position of the rehypothecation lender become much more riskier, in that only the long positions that had been rehypothecated were transferred. The rehypothecation lender had no other option but to sell the assets (an endogenous constraint emerged). Such an action (a best response in a game-theoretic sense) reverberated on the prime broker, that was forced as well to call funds. The consequence on the hedge fund was a sudden change in the duration of loans relative to the expected time to convergence of the arbitrage opportunities on the asset side of the balance sheet: "from an arbitrageur's perspective, seemingly long-term capital became truly short-term overnight" (ivi, p. 490). We refer the reader to the original paper for a systematic discussion of such phenomenology, which is our aim to enlighten in terms of first principles.

Let us try and sharpen the sense in which such phenomenology seems to display limited cognition about liquidity dynamics. On general grounds one may well agree that the evolution of market practices has been projecting market participant into "uncharted territory", whose cognition, therefore, can be reasonably expected to take time and effort. On somewhat more specific grounds, we seem to notice a conceptual gap between a scenario analysis, which may have anticipated the unfortunate chain of events, and a first principle analysis, which provides, so to say, high-level insights about the economic phenomena at play. By employing a somewhat philosophical jargon, we may call *extensional* a scenario analysis, meant to contemplate a large (complete?) set of scenarios in order to ascertain the potential emergence of distress under definite circumstances. Correspondingly, we may call *intensional* a first principle analysis, which aims at identifying the nature of a problem, irrespective of the circumstances which may characterize any empirical chain of facts. The relevance of first principles in the analysis of financial stability is by now quite established (see for instance Mehrling, 2012a). In essence, we advocate an intensional approach on liquidity cognition, which, for the case study under consideration, goes as follows.

The very fact that the financing mechanism *h* entails two steps (from the rehypothecation lender to the prime broker, and subsequently to the hedge fund) makes it evident that both steps need be performed in order for the process to work, and that such a complication, arguably, makes *h* 'riskier' with respect to a one step relation. In fact, the specific mechanisms which displayed the instability of *h* at the filing of Lehman are discussed by Mitchell and Pulvino (2012), and might have been anticipated by a far from straightforward systematic scenario (extensional cognition) analysis. On the other hand, a first principles analysis seems to provide a somewhat straightforward (intensional) cognition that *h does not embody the increasing strategic power which is embodied by the upper levels of H.* Upper levels of *h* do not represent higher strategic power, and are in fact subject to the same instabilities to which

lower levels are exposed, in particular concerning the value of collateral. In turn, instability in upper levels of *h* can transmit directly to levels below. As a result, what "seemingly" appeared to be stable (long-term capital) under normal conditions, turned out to manifest instabilities (mainly as transition to short-term capital).

Interesting further aspects relate to the fact that the strategic power embodied by a lending chain can be limited by the norms to which levels are subject. For instance, as reported by Mitchell and Pulvino (2012), in the US MMFs are required by the SEC to sell collateral in case of counterparty default (ivi, p. 473). Such a norm, meant to stabilize the system, limits the strategic power of a MMF under the circumstances discussed, and witnesses a deviation from *H*: the MMF cannot exert a disciplinary role, to the extent that its strategic power is limited (among other things) by norms.

Finally, abstracting from the specific case study, we seem to envision a line of inquiry about the limits of arbitrage connected with potential emergence of liquidity constraints, in particular with respect to arbitrages as activities – not confined to specific entities – which in fact contribute themselves to supply funding and market liquidity to the system (think of the role of dealers in both capital markets and money markets). The ability of arbitrageurs to perform countercyclical (stabilizing) activity seems highly dependent on such issues.

In such respects, the inherent hierarchy of money seems to provide a cogent standpoint for envisioning the emergence of endogenous liquidity constraints, with respect to which intensional cognition about liquidity risk – building on *H* – complements extensional cognition and standard scenario analysis or stress test. Needless to say, we do not claim to have provided even a partial solution to the complex liquidity problems discussed by BdF (2008) and Tirole (2011). Still, valuable insights seem to emerge in our behavioral perspective. Upper levels of *h* did not represent higher strategic power, and therefore did not embody the disciplinary role which underlies the stability of *H*. In our vision, it is such a disciplinary role which should guide one's vision (and consequently cognition) on the stability of credit chains (in which the role of collateral may be pivotal). Such matters do not seem to have been fixed as yet: according to the Financial Stability Oversight Council (2016), "The potential for fire sales of collateral by creditors of a defaulted broker-dealer also remains a significant risk."

6 Conclusions

On both theoretical and empirical grounds, a consensus seems to emerge about the pivotal role of the actual microstructure of financial markets in driving actual phenomena, in particular, liquidity dynamics at the systemic level. In the words of Caruana and Kodres, "liquidity is created and maintained by the market participants themselves" (BdF, 2008, p. 65). In fact, in recent years, the profound connection between liquidity, speed of capital and arbitrage activity has been attracting increasing interest (see for instance Mitchell et al., 2007; Gromb and Vayanos, 2010; Duffie, 2010; Teo,

2011). In our view, the inherent hierarchy of money provides a robust pillar for enlarging such a line of inquiry.

We have been arguing about enlarging the analysis of the limits of arbitrage towards cognition of liquidity conditions. Our theoretical proposal is to widen the reach of behavioural finance so as to embrace the interface between financial and monetary phenomena, which nowadays represents a major challenge for both academics and policymakers. The inherent hierarchy of money seems to provide a cogent perspective on the limits of cognition about liquidity and stability conditions. The limits of professional arbitrage thereby envisioned stem from unanticipated emergence of endogenous constraints on the strategic space of the arbitrageur's counterparties, which, in turn, reverberate on the arbitrageurs' behavior.

Such an approach may shed further light on the criticalities of the current financial outlook. Still, in our view, it is the theoretical relevance of the analysis of arbitrage which one should focus first, since "Modern finance theory rests on the ability of arbitrageurs to ensure that substantially similar assets trade at substantially similar prices" (Mitchell and Pulvino, 2012). The 'limits' of financial models have been long pointed out on general grounds; such models are not meant to represent the "laws" governing the dynamics of financial markets, which are far more complex than physical systems subject to well defined laws of motion. Such models represent benchmark analytic frameworks in which no arbitrage conditions embody assumptions of perfect liquidity. True, on somewhat more specific grounds, our behavioral approach to the limits of arbitrage may help sharpen the interpretation of risk models, whose current appreciation has been challenged by a number of scholars: in the words of De Bondt et al. (2008), "quantitative risk models disregard rare events and try to model what arguably cannot be modeled" (ivi, p. 17).

The analytical challenge of embedding the speed of capital and the limits of arbitrage into a theory of asset pricing may not be unsurmountable (see for instance Duffie, 2010), and may contribute to fix a shared cognition about the properties of liquidity dynamics. According to William C. Dudley,[9] "We need to better understand the degree to which any changes in the nature of liquidity reflect the evolving structure of financial markets, changes in regulatory power or other factors." Hopefully, our insights on liquidity cognition and limits of arbitrage may prove useful in such respects. One may argue that a game-theoretic approach may embrace both Shubik's (1999) strategic approach to monetary phenomena and Walliser's (2008) framework for cognition and coordination. Perhaps the issue can be raised of whether arbitrage and liquidity can be considered, in some sense, as sides of the same coin.

A fruitful line of progress of behavioral finance seems to emerge, in connection with potential applications of our approach to the debate about the stabilization of markets. We are still experiencing the consequences of the Great Recession, with its tragic impact on labor demand and welfare levels. It

[9] Federal Reserve Bank of New York, "Regulation and liquidity provision." Speech, September 30, 2015.

is quite natural to argue that less constrained and more stable arbitrage activity may help stabilize the financial system; it is far more engaging to imagine that progresses in the understanding of the limits of arbitrage may help envision "ways" of reducing substantially the impact of financial downturns on income and wealth distributions. Shiller (2014) advocates the emergence of markets for claims on the flows of gross domestic products and other macroeconomic aggregates, arguing that such markets would help countries share their risks. At a micro level as well, according the Nobel laureate, financial innovation may help improve the stability of income distributions (people hedging themselves in occupational markets). Along such perspectives, hopefully, our approach to the limits of arbitrage may turn out to define sound building blocks of analysis.

References

Acharya, V. V., Cooley, T., Richardson, M., Walter, I. (2009). Manufacturing tail risk: a perspective on the financial crisis of 2007-2009. Foundations and Trends in Finance 4(4): 247-325.

Araújo, A., Schommer, S., Woodford, M. (2015). Conventional and unconventional monetary policy with endogenous collateral constrains. AEJ: Macroeconomics 7(1): 1-43.

Barberis, N., Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance. Elsevier Science B.V.

Bell, S. (2001). The role of the state and the hierarchy of money. Cambridge Journal of Economics 25: 149-163.

BdF (Banque de France, 2008). Financial Stability Review - Special Issue. Liquidity. February.

Bengtsson, E. (2013). Shadow banking and financial stability: European money market funds in the global financial crisis. Journal of International Money and Finance 32: 579-594.

Bhanot, K., Guo, L. (2012). Types of liquidity and limits to arbitrage – the case of credit default swaps. Journal of Futures Markets 32(4): 301-329.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch of 2007-2008. Journal of Economic Perspectives 23(1): 77-100.

Brunnermeier, M. K., Pedersen, L. H. (2009). Market liquidity and funding liquidity. Review of Financial Studies 22(6): 2201-2238.

Brunnermeier, M. K., Sannikov, Y. (2016). The I theory of money. Unpublished.

Cochrane, J. H. (2005). Asset Pricing. Revised edition. Princeton University Press.

De Bondt, W., Muradoglu, G., Shefrin, H., Staikouras, S. K. (2008). Behavioral finance: quo vadis? Journal of Applied Finance – Fall/Winter: 7-21.

Dittmar, A. (2008). Corporate cash policy and how to manage it with stock repurchases. Journal of Applied Corporate Finance 20(3): 22-34.

Duffie, D. (2010). Asset price dynamics with slow moving capital. Journal of Finance LXV(4): 1237-1269.

Financial Stability Oversight Council (2016). Annual Report.

Grillet-Aubert, L., Haquin, J. B., Jackson, C., Killeen, N., Weistroffer, C. (2016). Assessing shadow banking – non-bank financial intermediation in Europe. European Systemic Risk Board.

Gromb, D., Vayanos, D. (2010). A model of financial market liquidity based on intermediary capital. Journal of the European Economic Association 8(2-3): 456-466.

Holmström, B., Tirole, J. (2011). Inside and outside liquidity. The MIT Press.

Hull, J. C. (2007). Risk management and financial institutions. Pearson.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. American Economic Review 93(5): 1449-1475.

MacKenzie, D. (2006). An engine, not a camera. The MIT Press.

Mehrling, P. (2012a). Three principles for market-based credit regulation. American Economic Review 102: 107-112.

Mehrling, P. (2012b). The inherent hierarchy of money. Social Fairness and Economics: Essays in the Spirit of Duncan Foley.

Mitchell, M., Pedersen, L. H., Pulvino, T. (2007). Slow moving capital. American Economic Review 97: 215-220.

Mitchell, M., Pulvino, T. (2012). Arbitrage crashes and the speed of capital. Journal of Financial Economics 104: 469-490.

Nashikkar, A., Subrahmanyam, M. G., Mahanti, S. (2011). Liquidity and arbitrage in the market for credit risk. Journal of Financial and Quantitative Analysis 46(3): 627-656.

Pozsar, Z. (2014). Shadow banking: the money view. Office of Financial Research.

Shiller, R. J. (2014). Speculative asset prices. American Economic Review 104(6): 1486-1517.

Shleifer, A., Vishny, R. W. (1997). The limits of arbitrage. Journal of Finance LII(1): 35-55.

Shleifer, A. (2000). Inefficient markets. An introduction to behavioral finance. Clarendon Lectures in Economics. Oxford University Press.

Shubik, M. (1999). The theory of money and financial institutions. Vol. 1. The MIT Press.

Shubik, M. (2010). The theory of money and financial institutions. Vol. 3. The MIT Press.

Singh, M. (2013). The changing collateral space. IMF Working Paper 13/25.

Soros, G. (2013). Fallibility, reflexivity and the human uncertainty principle. Journal of Economic Methodology 20(4): 309-329.

Stiglitz, J., 2012. Macroeconomics, monetary policy, and the crisis. In the Wake of the Crisis, Blanchard, O. et al. eds. MIT Press.

Stiglitz, J. (2015). Reconstructing macroeconomic theory to manage economic policy. Fruitful Economics. Essays in honor of and by Jean Paul Fitoussi. Laurent, E., and Le Cacheux, J. eds.

Teo, M. (2011). The liquidity risk of liquid hedge funds. Journal of Financial Economics 100: 24-44.

The Economist (2008). Confessions of a risk manager. August 7th.

The Economist (2016). Minsky's moment. July 30th.

Tirole, J. (2011). Illiquidity and all its friends. Journal of Economic Literature 49(2); 287-325.

Walliser, B. (2008). Cognitive economics. Springer.

Xu, L., Yin, X., (2017). Exchange traded funds and stock market volatility. International Review of Finance, doi:10.1111/irfi.12121.