SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di / Department of

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa

Dottorato di Ricerca in / PhD program Economia e Finanza Ciclo / Cycle XXXII

Essays on Financial Intermediation

Cognome / Surname Montoli Nome / Name Stefano

Matricola / Registration number 821689

Tutore / Tutor: Prof. Giovanna Iannantuoni

Supervisor: Prof. Vittoria Cerasi

Coordinatore / Coordinator: Prof. Matteo Manera

ANNO ACCADEMICO / ACADEMIC YEAR 2018/2019

Introduction

This thesis consists of two independent essays investigating different topics in financial intermediation.

The first chapter (co-authored with Vittoria Cerasi) analyses the impact of the resolution policy chosen by a (unique) regulator on the incentive of banks to expand abroad. In line with recent empirical evidence, we propose a simple model to capture the increased risk-sensitivity of banks' unsecured debt following the introduction of bail-in as the main resolution policy. Focusing on the *endogenously* determined cost of funding, we characterize the increased incentive for banks to increase geographic diversification. Searching for new lending opportunities in foreign countries, banks can counterbalance the greater funding cost with a reduction in credit risk. Building on this result, we complement the analysis addressing the optimal policy chosen by the regulator, whose objective is to maximize social welfare. In choosing the allocation of the burden of banks' failure between taxpayers and bondholders, the regulator faces a trade-off. In particular, she must weight the loss of welfare entailed by the drop off the market of the least efficient banks, due to an increase in funding costs, with their substitution with more efficient and diversified banks. As a result, we identify that a positive level of public support is optimal. Furthermore, the calibration of the optimal resolution policy depends on the social value of banking services and on the soundness of the banking sector.

The second chapter (co-authored with Davide Bosco) explores the determinants of financial fragility for non bank-financial intermediaries. In particular we study the coordination problem for the investors of an open-end financial institution. Reacting to the (private) observation of bad news - possibly due to a temporary negative shock - investors might opt for an early redemption of their shares. To respond to the liquidity needs, the fund starts selling some of its assets. In doing so, it signals to the market that its fundamentals may be bad. The consequent reduction in potential buyers' willingness to pay deflates asset prices, thus forcing the institution into a fire sales spiral. We show that, even within the framework of informationally efficient markets - to some extent, *because* of the informational efficiency - liquidity needs might lead to a balance sheet degradation cycle that could potentially push a sound financial institutions into insolvency. This results posits that the well documented causal link between insolvency and illiquidity can work also in the opposite direction. Furthermore, we comment on the nature of the strategic interaction between the fund's investors. Strategic complementarity arises endogenously when investors receive sufficiently bad information.

Bank resolution and multinational banks

Vittoria Cerasi* Stefano Montoli †

November 2020

Abstract

This paper studies the impact of different resolution policies on banks' choice to expand abroad. The regulator can choose to resolve banks through *bail-in* or *bail-out* or a combination of the two. The choice of the regulator affects the cost of funding of banks, endogenous in the model. We study the relative profitability of alternative bank corporate structures, either multinational (large and diversified) or domestic (small and non diversified) for different levels of public support. Our model allows us to identify the potential impact of the resolution policy on the structure of the banking system. Lower levels of public support increase the cost of funding for all banks, in line with recent empirical evidence. We show that a reduction in the level of public support (from bail-out to bail-in) induces banks to expand abroad in search for alternatives to save on their funding costs. Finally, we are able to identify the optimal resolution mix by taking into account the reaction of banks to the policy.

JEL classification: G21, G28.

Keywords: Bank regulation; bail-in; multinational banks; bank funding cost.

*Bicocca University, Department of Economics, Management and Statistics (DEMS) and Center for European Studies (CefES), Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy; e-mail: vittoria.cerasi@unimib.it

†Bicocca University and Center for European Studies (CefES). Email: s.montoli3@campus.unimib.it.

Acknowledgements. We thank Nicola Limodio, Xavier Freixas and Giacomo Calzolari for their comments on a preliminary version of this paper.

1. Introduction

The introduction of bail-in has been one of the major changes in banking regulation in the aftermath of the Global Financial Crisis. It is widely acknowledged that the new regulatory framework represents a significant improvement in the resolution of banks. However, a comprehensive assessment of the impact of this new regulation requires a better understanding of how banks will individually react to this new policy.

Recent empirical evidence has documented an increase in banks funding costs following the adoption of the new resolution regime. We argue that individual banks could seek to reduce their funding costs by increasing their geographic diversification, thus reducing their exposure to credit risk. In particular, this paper proposes a simple framework to capture the impact of bail-in on funding costs and its implication for banks' incentive to expand abroad. The final goal is to identify the optimal resolution policy, taking into account the reaction of banks to the resolution framework.

The global financial crisis has shown how costly it can be to rely on public interventions to preserve financial stability and support troubled banks. Indeed, while governments' support has guaranteed continuity in the provision of banking services, it has led to unprecedented costs for public finances in many developed countries. In Europe, the Commission authorized total aids of EUR 3,892.6 billion for guarantees on liabilities between 2008 and 2014. An additional EUR 448 billion was spent on the public recapitalization of banks between 2008 and 2013 (Lintner *et al.*, 2016). In the US, after the failure of Lehman Brothers, several financial institutions received public support, mainly through the Troubled Asset Relief Programme, which accounted for 6% of the GDP in the fourth quarter of 2008 (Philippon and Salord, 2017). As a result, many countries, especially in Europe, experienced a severe sovereign debt crisis in 2010-12. The vicious circle, the so called "doom loop", between a rising sovereign debt and banking crises has been recently analyzed in Acharya, Drechsler, and Schnabl (2014). In addition to absorbing public resources, bailouts are also lined to distortions of market incentives: the anticipation of ex-post public support in case of distress may weaken market discipline, leading ex-ante to excessive risk taking from banks[1].
As a solution to the outlined problems, most developed economies have introduced formal bank resolutions and bail-in regimes (U.S. Dodd-Frank Act or the European Bank Recovery and Resolution Directive (BRRD)). According to this new regulation, in case of banks' distress creditors will directly absorb banks' losses, thus bearing part of the costs of restoring the bank. The BRRD allows for public support to failing institutions, when this is required to preserve financial stability, but only after shareholders and creditors have contributed

[1]See Dell'Ariccia *et al.* (2018) and the references therein for empirical evidence.

with their own funds up to 8% of the bank liabilities. The objective of the new regulatory framework is to minimize the cost of banks recovery for taxpayers and to improve the efficiency of the resolution process. Also, turning unsecured debt into bailinable debt should increase the level of monitoring by creditors, thereby reducing banks' risk-shifting attitude.

Recent empirical evidence, however, has documented an increase in banks' funding costs, as creditors become mere exposed to losses. In particular, Schafer, Schnabel, and Weder (2016) have shown that share price movements signal that the expectation of bail-in indirectly affects future bank returns, through the effect on [increased] funding costs. Crespi, Giacomini, and Mascia (2019), Cutura (2018), Giuliana (2019), Lewrick, Garralda, and Turner (2019) document an increase in spread between bailinable and non bailinable bonds for European banks. Furthermore, the spreads appear to be sensitive to banks' riskiness, thus supporting the hypothesis that bail-in induced greater market discipline.

We elaborate on the strategic reaction of banks to the new resolution policy, by focusing on the increased sensitivity of funding costs to banks' riskiness. In particular, we unveil a relation between bail-in and geographic diversification. Indeed, while the bail-in regime induces investors to require a risk premium for holding bank bonds, risk can be diversified away by banks through foreign expansion. Levine, Lin, and Xie (2019) demonstrate that geographic diversification is linked to lower costs for interest bearing liabilities. Aldasoro, Hardy, and Jager (2020) also support the hypothesis that geographic complexity reduce banks' exposure to local shocks, thanks to an increased diversification.

Our paper provides a theoretical framework to analyze the link between resolution policies and banks' decision to expand abroad. In our model, banks are characterized by heterogeneous cost of monitoring the projects they finance and funding costs are determined endogenously. Banks can choose to operate domestically, lending in one single country, or as multinational bank, lending in two countries. Multinational banks are bigger and, thanks to the diversification of their portfolio of loans, can allocate resources across units to reduce their exposure to credit risk. Indeed, credit losses can be shared across branches and resources from the branch operating in one country can be used to cover potential shortfalls in the other country.

In case banks are not able to meet their obligations, due to large credit losses, a group of bondholders may be subject to haircut on the face value of their bonds, while the remaining group is reimbursed using public money. In the model, we allow for a resolution policy ranging from full bail-out to full bail-in. The focus of our analysis is to understand how the resolution policy impacts on the endogenous funding cost of banks.

The novelty of our paper is to show that bail-in, not only induces a rise in the overall funding cost for banks, but it also stimulates banks to expand abroad in search for gains from diversification. This implies that moving from bail-out to bail-in, the number of viable domestic banks shrinks, although this effect can be more than compensated by the entry of new multinational banks.

The key mechanism is the following. When there is a positive probability of bailout, the risk entailed in a portfolio of loans in a bank is not perfectly priced, due to the implicit guarantee of public subsidies. This creates a wedge in expected funding costs between domestic and multinational banks. In other words, the choice to expand abroad reduces the benefits of public guarantees. This is a perverse effect of diversification for multinational banks. Conversely, domestic banks fully benefit from bailout, as they maximize the likelihood of receiving public support. When the resolution regime moves from bailout to bail-in, bonds' prices start to correctly reflect the true riskiness of banks. As such, multinational banks enjoy lower funding costs due to the diversification of credit risks. To conclude, when we rule out public support, multinational banks benefit from a reduced risk premium on bail-in-able debt, thus becoming relatively more profitable.

In determining the optimal policy, the regulator anticipates the reaction of the banking sector. As such, the trade-off that allows for the identification of the optimal level of public support is the following: a reduction in the level of public support increases bank funding cost for all bank structures, thus increasing the share of insolvent domestic banks. On the other hand, it increases the relative profitability of multinational banks. The social benefit of multinational banks is twofold: first, the supply of credit is expanded due to the substitution of insolvent domestic banks with larger multinational banks. Second, thanks to their diversification, multinational banks are less risky, although they involve a greater cost of supervision due to their complexity. The balance between the social costs of increased failures of domestic banks and the benefits from a higher share of multinational banks allows us to identify the optimal resolution regime (i.e. the optimal level of public support).

Relation with the literature. Our paper contributes to the literature on the optimal design of banks' resolution policy. Within the intense debate over the reforms that followed the Great Financial Crisis, a strand of literature has focused on the pros and cons of the different policies to deal with troubled banks. In particular, Avgouleas and Goodhart (2015) argue that, while bail-in induces better creditors monitoring and reduces the exposure to moral hazard, it also leads to higher funding costs for banks and might amplify banks' crisis in case of systemic shocks. In a similar vein, Dewatripont (2014) argues that bail-in can be detrimental to financial stability. Indeed it might increase the bank runs, especially if banks don't hold enough long-term loss absorbency capacity to prevent panic from spreading among short-term claim-holders. Moreover, bailout might help protecting the "average bank" in case of a systemic crisis, leaving only the worst banks out of the market. This observation is in line with the optimal policy proposed in our analysis. Indeed, we argue that bail-in creates an incentive for more efficient banks to expand abroad, while the least efficient banks are forced to close. Nevertheless, a certain level of public support is desirable to prevent the failure of a sub optimally large number of banks. Dell'Ariccia *et al.* (2018) review the empirical literature on the drawbacks of the two resolution regimes: the nexus between bailout and banks' moral hazard and the systemic spillovers of bail-in. Furthermore, the authors introduce a theoretical model to account for the costs associated with bail-in and bail-out, showing that the optimal policy mix crucially depends on the relevance of systemic spillovers of bail-in. Although all these papers emphasize the benefits of bail-in, they argue in favour of maintaining some degree of public intervention, especially when needed to preserve financial stability. In line with this, we show that, considering the impact of the resolution policy on the structure of the banking sector, some public support is necessary to avoid the exit from the market of an excessively large share of banks.

Within the debate on the optimal mix between bail-out and bail-in, Walther and White (2019) and Pandolfi (2018) show that it is optimal to complement bail-in with bail-out to reduce deadweight social losses. While in Walther and White (2019) bail-in consists in writing-off long-term debt and ameliorates the debt-overhang problem, the discretion left to the regulator to call for bail-in releases a bad signal to investors who refuse to roll-over short-term debt. To avoid an aggressive bail-in for banks that do not deserve it in some states of the world, they call for some degree of bail-out. In Pandolfi (2018) bail-in is defined as conversion of debt into equity and it implies dilution of incentives to monitor loans for the insiders of the bank. In addition, the higher funding cost required by bondholders as a consequence of expectation of bail-in, reinforces the debt-overhang problem of the banker. To avoid this, bailout could be retrieved in some instances. Both papers analyze the consequences of bail-in on the cost of funding for banks and on the incentives of the banker, calling for a mix of bail-in and bail-out. We also elaborate on the endogenous rise of funding costs induced by the bail-in and solve for the optimal resolution policy mix. However we focus on the diversification choice of the banker and show that bail-in is the optimal solution when considering the impact on the structure of the banking industry, although it implies exit of the least efficient banks. If this involves a social cost due to the disruption of payment services, there is scope to restore some degree of bailout.

The nexus between the resolution policy and banks' choice to expand abroad, in presence of endogenously determined funding costs, is also studied in Luciano and Wihlborg (2018). While allowing for a richer characterization of banks' organizational structure (e.g. different forms of foreign of representation are considered), their focus lies on the value of the different organizational structures and their implications for systemic risk. The different objectives of the banks (the maximization of value) and the regulator (the minimization of systemic risk) gives rise to a conflict of interest between the two parties. In this sense, our analysis is complementary, in the sense that it enriches the interaction between the banks and a benevolent regulator. Indeed, our simpler framework allows for the aggregation of banks' organizational choices, which in turn is key for the identification of an optimal resolution policy.

Our analysis also builds on the empirical evidence of the adoption of the new resolution policy on bank funding costs. Schfer *et al.* (2016) document a decrease in bank returns following a bail-in. The authors interpret this evidence as suggestive of an expected increase in funding costs due to the reduced likelihood of future public support. Crespi, Giacomini and Mascia (2019) find evidence of an increase in the issuance spreads between bail-in-able bonds and non bail-in-able bonds for Italian banks. This effect became significantly relevant after Italian authorities decided to resolve four small banks in November 2015. Cutura (2018) and Giuliana (2019) also find similar evidence, finding an average 10 basis points bail-in premium for bail-in-able bonds. Similarly, Lewrick, Garralda and Turner (2019) identify a bail-in risk premium for bail-in-able bonds issued globally. This literature documents how the introduction of the bail-in regime was deemed credible by investors and how it increased banks' market discipline. Indeed, the increase in funding costs also appears to depend on banks' riskiness. As such, this evidence also support our result that a switch from bail-in to bail-out increases the sensitivity of funding costs to the risk of the banks' portfolios.
Our work also relates to the evidence in Levine, Lin and Xie (2019), which shows that geographic expansion reduces the costs of interest-bearing liabilities thanks to the benefits of diversification. This evidence is further supported by the observation that the drop in funding costs is more pronounced when the bank expands in regions whose economy is less correlated to the region of origin of the bank. Further studies directly addressed the impact of banks' geographic expansion on their riskiness. Deng and Elyasiani (2008) and Goetz, Laeven and Levine (2016) analyse geographic diversification in the US banking industry, showing that geographic expansion lowers risk thanks to a reduced exposure to idiosyncratic shocks. Fang and van Lelyveld (2014) use a sample of some of the world's largest banking groups to show that geographic diversification helps mitigating credit risk, although the magnitude of the effect varies significantly across different banking groups. In particular the benefits of diversification appear to be dependent on the level of synchronization of the business cycle in different countries. Finally, Faia, Ottaviano and Sanchez (2017) and Duijm and Schoenmaker (2020) provide empirical evidence on the negative link between geographic diversification and riskiness of European banks. They confirm that diversification is beneficial also when international rather than intra-national, and more so when business cycles of the

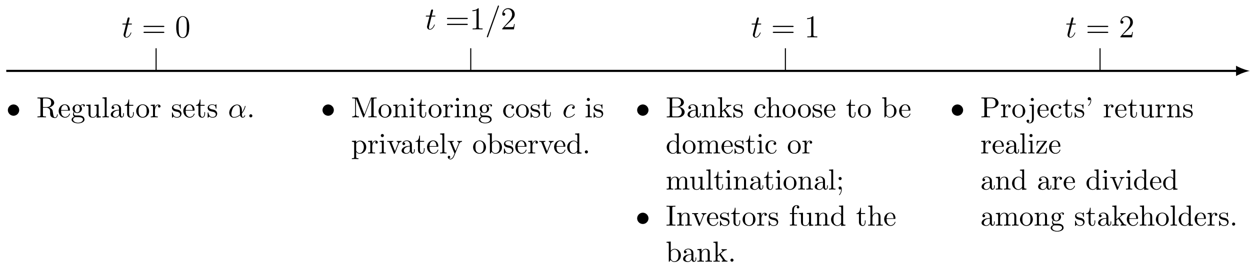

Fig. 1. Timeline of the game

different countries where the bank is located are less synchronized. Overall, this literature provides support to our observation that banks' decision to expand abroad is beneficial to their risk profile thanks to an increased resilience to idiosyncratic shocks, and this in turn is reflected on banks' funding costs.

The rest of the paper is organized as follows: section 2 sets up the model, section 3 describes the effect of different levels of public support, section 4 analyzes the optimal resolution policy and section 5 concludes.

2. The Model

Consider an economy populated by banks, entrepreneurs with productive projects and investors. All agents are risk neutral.

Projects. Each productive project yields a return R with probability $p \in (0.5, 1)$, 0 otherwise.

Banks. Banks have no capital and finance risky projects by issuing bonds. Banks monitor projects at a cost c. There is heterogeneity at the bank level, that is, each bank is characterized by a specific cost of monitoring c distributed according to a uniform distribution on $[0, c_{Max}]$.

Investors. There is an infinite number of investors with one unit of capital each. Investors can choose between bonds issued by banks and an alternative investment yielding a safe return equal to 1. Capital markets are perfectly competitive.

Resolution policy. Assume bondholders might benefit of a public support in the form of a transfer in case of bank failure: a fraction $\alpha \in [0, 1]$ of bondholders is protected, while the remaining fraction suffer losses in case of bank failure. As α increases, we move from the case of complete bail-in ($\alpha = 0$) to the case of full bail-out ($\alpha = 1$).

Timing. The timing of the game is the following: first the regulator announces the fraction of bondholders that will be compensated with public money in case of bank resolution, that is α; then, each bank observes its specific realization of the cost of monitoring c; investors set the rate at which they are ready to fund each bank. Finally, returns of the lending activity realize and stakeholders are repaid. The timing of the game is represented in Figure 1. The game is solved backward.

2.1. Endogenous funding cost

The objective is to determine endogenously the cost of funding for the bank when there is a resolution policy in place. For simplicity we assume that the bank retains the whole return R from the productive projects it finances. In this case, the expected profit of the bank is:

$$p \times max\{R - r_D, 0\} - \frac{c}{2} \tag{1}$$

From equation (1) we see that *ex-ante* the bank is solvent whenever $R > r_D$, that is, only in case the return of the successful project is sufficiently high to repay bondholders. In case of failure of the project, due to limited liability, the bank is insolvent and does not repay bondholders. We assume that each project has a positive NPV, that is:

$$R > \frac{1}{p} > 1 \tag{A1}$$

with the limitation that $p \in (0.5, 1)$, namely $\frac{1}{p} < 2$.

Due to competition in capital markets, investors expect to receive just the opportunity cost of their initial investment, set equal to the return on the safe asset. Investors anticipate to be repaid the face value r_D when the bank is solvent; when the bank is insolvent, since investors observe the resolution policy, they anticipate that a fraction α of bondholders will recover at least the opportunity cost of their investment through injection of public money.

The investors' rationality condition is:

$$pr_D + (1 - p)\alpha = r_D - ES_1(\alpha) = 1 \tag{2}$$

where $ES_1(\alpha) \equiv (1 - p)(r_D - \alpha)$ are the *ex-post* expected shortfalls that bondholders anticipate to suffer. In equilibrium, the cost of funding for a domestic bank is:

$$r_D(\alpha) = \frac{1 - \alpha}{p} + \alpha \tag{3}$$

Notice that the cost of funding decreases with α and reaches its minimum, i.e. $r_D = 1$, when there is full bail-out ($\alpha = 1$). As a matter of fact, the derivative of (3) w.r.t. α is the "odd ratio" of the event default, namely the probability that the domestic bank does not repay bondholders over the probability that it repays them:

$$\frac{\partial r_D(\alpha)}{\partial \alpha} = -\frac{(1-p)}{p} < 0 \tag{4}$$

As the expectation of bail-out increases, the interest rate required by bondholders falls, as the probability of having to face *ex-post* expected shortfalls reduces. This result is in line with the empirical evidence finding that after the introduction of bail-in the cost of funding has risen for banks as more creditors have started to expect an increasing probability of suffering an haircut on the face value of their bonds in case of bank failure. In case of full bail-out bondholders are ready to accept the minimum interest rate, that is, the return on perfectly insured deposits.

Substituting r_D into (1), we derive the expected profit of the bank at the equilibrium:

$$E[\Pi_D(\alpha)] = pR - 1 + \alpha(1-p) - \frac{c}{2} \tag{5}$$

where the first term represents the NPV of the productive project, the second is the amount of public money injected in case of bank failure and the third is the cost of monitoring. The bank pays r_D every time the project succeeds. As α increases r_D decreases, implying savings for the bank on the amount of money owned to its creditors. With bail-out the public money replaces the private money with which the bank has to repay bondholders, reducing the cost of funding for the bank. These savings become an extra-profit $\alpha(1-p)$ for the bank.

The bank is viable, that is, its profit is non-negative, when the monitoring cost is sufficiently low:

$$c \leq c_D(\alpha) \equiv 2(pR - 1) + 2\alpha(1-p) \tag{6}$$

Notice that the threshold $c_D(\alpha)$ is increasing in α since $c'_D(\alpha) = 2(1-p) > 0$. Indeed, greater public support, increases the slack, inducing even less efficient banks to become viable.

2.2. Multinational banks

Consider now the strategy of a bank willing to increase its scale of lending. Assume that the bank faces no opportunities in the home country and thus has necessarily to expand abroad. In our simple model we imagine a bank financing two projects, one project in the home country and the other in the foreign country. Although returning the same return R the two projects are uncorrelated, thus giving rise to benefits from diversification. Assume that all banks in the economy face this choice between operating in the home country, i.e. to be *domestic*, or expanding abroad, thus becoming a *multi-national* bank (MNB).

When a bank chooses to be domestic it raises one unit of debt and finances one project in the home country. The alternative is to become a MNB, thus raising two units of debt to finance two projects, one in each country. Expanding abroad means opening a branch, that is, a foreign office of the home bank[2]. To capture the idea of an increasing cost of monitoring uncorrelated projects and more in general the costs of running two units in two different countries, we assume that MNBs face twice the cost of monitoring one project by a domestic bank, since lending abroad implies overcoming legal and cultural barriers.

Considering that both the domestic and the foreign units are responsible for each other's losses, we can write the consolidated profits as:

$$2p^2 \times max\{R - r_M, 0\} + 2p(1-p) \times max\{R - 2r_M, 0\} - c \tag{7}$$

The central term in the profit refers to the case where one project is successful and the other one fails (with probability $p(1-p)$). We thus have to distinguish between two cases:

- case (a), when $r_M \leq \frac{R}{2}$: the return from the successful project is enough to repay the promised rate to bondholders;
- case (b), when $r_M > \frac{R}{2}$: what is returned by the successful project is not enough to refund all bondholders; in this case $(1-\alpha)$ bondholders receive $\frac{R}{2}$, while another fraction α is bailed-out and receive 1 unit as public support.

In the rest of the paper we will develop the analysis assuming

$$R \geq \frac{2}{p(2-p)} \tag{A2}$$

which refers to case (a). In particular, we assume that the returns on the financed projects are sufficiently high that multinational banks can always repay both depositors even if only one project is successful. This holds independently from the resolution policy chosen by the regulator.

The investors' participation constraint is:

$$p^2 r_M + 2p(1-p)r_M + (1-p)^2\alpha = 1 \tag{8}$$

The first term is when both projects succeed with probability p^2 and bondholders receive the promised rate r_M; the second term is when one project succeeds and the other fails, but the

[2] The case where a bank decides to expand abroad *via* a subsidiary in analysed in appendix A.

revenue $R/2$ is enough to repay r_M to each bondholders; the last term is when both projects fail with probability $(1-p)^2$ and a fraction α of bondholders is bailed out and receive 1 unit of public money.

The expected shortfalls for bondholders are:

$$ES_2^a(\alpha) = (1-p)^2(r_M - \alpha) \tag{9}$$

since bondholders expect to suffer *ex-post* losses equal to $(r_M - \alpha)$ in case the MNB fails. The overall cost of funding for a MNB, solving equation (8) for r_M, can thus be written as:

$$r_M^a(\alpha) = \frac{1-\alpha}{p(2-p)} + \alpha \tag{10}$$

Notice that the cost of funding decreases with α and reaches its minimum, i.e. $r_M^a = 1$, when there is full bail-out ($\alpha = 1$). As a matter of fact, the derivative of (10) w.r.t. α is the "odd ratio" of the event default, namely the probability that the MNB does not repay bondholders on the probability that it repays them:

$$\frac{\partial r_M^a(\alpha)}{\partial \alpha} = -\frac{(1-p)^2}{1-(1-p)^2} < 0 \tag{11}$$

As the expectation of bail-out increases, the interest rate required by bondholders falls. The probability of having to face a default *ex-post* reduces with α, similarly to the case of a domestic bank. However an increase in bail-out reduces the odd ratio **more** for a domestic bank compared to a MNB. This is due to the fact that, with some degree of bail-in, the probability that a domestic bank defaults is **larger** compared to a MNB, since a MNB gains from diversification of its portfolio of loans. In other words, bail-out benefits **more** bondholders of a domestic bank compared to those of a MNB, since the MNB repays them more often with the greater expected revenue from its diversified portfolio of loans.
We can now write the expected profits of the bank, substituting the bond rate (10) into (7):

$$E[\Pi_M^a(\alpha)] = c_D(\alpha) - 2\alpha p(1-p) - c \tag{12}$$

using the definition of the threshold for the domestic bank in (6).
Based on what derived so far, we have that a MNB is profitable only when the monitoring cost is below a specific threshold, that is:

$$c \leq c_M(\alpha) \equiv c_D(\alpha) - 2\alpha p(1-p) \tag{13}$$

It is immediate to see that $c_M(\alpha)$ is smaller than $c_D(\alpha)$: a MNB must be more efficient compared to a domestic bank in order to be viable. This is because the domestic bank maximizes the expected value of public subsidies. Indeed, the revenues and the cost of funding excluding public intervention are equal for the two bank structures, while the expected value of the public subsidies is different. In particular, the difference between c_M and c_D is equal to the difference in the expected value of public interventions for the two bank structures. Notice that also for a MNB the threshold to be viable increases with α.

3. Resolution and banking structure

Each bank is identified by a specific monitoring cost c. According to the region where this specific c lies, the bank may be viable as a domestic or multinational bank. However the relative profitability of the two type of banks is affected by the particular resolution policy, in our simple model defined by α. In this section we analyze how the resolution policy changes the structure of the banking system, namely the fraction of MNBs over domestic banks. We first consider the case of full bail-out ($\alpha = 1$) and then look at the case of complete bail-in ($\alpha = 0$).

3.1. Full bail-out ($\alpha = 1$)

In this section we consider the case where all investors are protected by public guarantees in case of bank failure, that is the case with $\alpha = 1$. In this case the bond rate is equal to 1 for both types of bank, as $r_D(1) = r_M^a(1) = 1$. Exploiting equation (6) when $\alpha = 1$, the domestic bank is viable when:

$$c \leq c_D(1) \equiv 2(pR - 1) + 2(1 - p) \tag{14}$$

while from equation (13) when $\alpha = 1$, a MNB is viable whenever:

$$c \leq c_M(1) \equiv c_D(1) - 2p(1 - p) < c_D(1) \tag{15}$$

In addition we derive the condition for which a MNB is more profitable than a domestic bank:

$$c \leq c_\Pi(1) \equiv c_D(1) - 4p(1 - p) < c_M(1) \tag{16}$$

In the case of full bail-out, we can summarize the results in the following proposition:

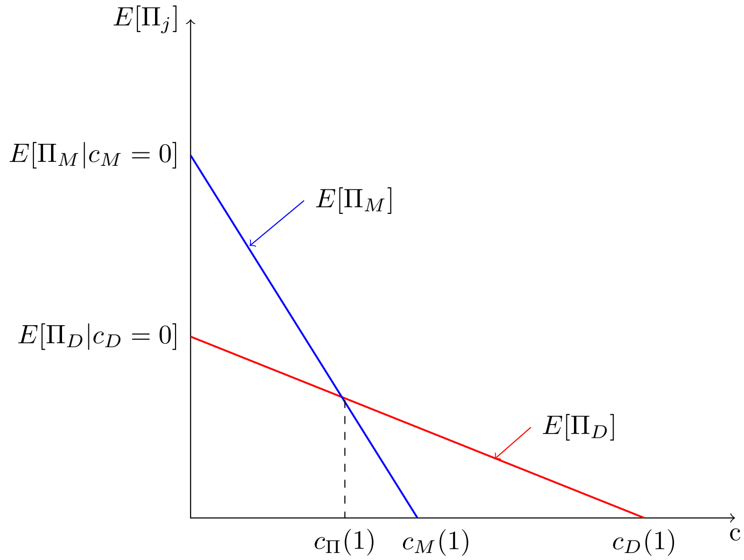

Fig. 2. Expected profits with $\alpha = 1$.

Proposition 1. *Assuming that each bank's monitoring cost is a specific realization from the uniform distribution on* $[0, c_{Max}]$*, projects' return is* $R \geq \frac{2}{p(2-p)}$ *with probability of success* $p \in (0.5, 1)$ *and returns are independent across countries, we can charaterize the structure of the banking sector in case of full bail-out* ($\alpha = 1$) *as follows:*

(i) Given $c_M(1) < c_D(1)$*, MNBs need to be more efficient than domestic banks in order to be viable. When* $c_M(1) < c < c_D(1)$ *domestic banks are viable, whereas MNBs are not;*

(ii) There exists a subset of realizations of c *where both MNBs and domestic banks are viable, but the latter is more profitable, that is* $c_\Pi(1) < c < c_M(1)$*;*

(iii) If $c < c_\Pi(1)$*, the expected profit of MNBs are higher compared to that of domestic banks.*

Results at points (*i*) and (*ii*) are explained by the fact that, for each unit lent, the domestic bank saves on funding costs since it pays the same face value (=safe rate) to its bondholders, but with a lower probability (p instead of $p(2-p)$ for a MNB): hence it can stands a lower degree of efficiency in monitoring. Result at point (*iii*) is explained by the smaller size of a domestic bank compared to a MNB: for low enough monitoring cost, the MNB has expected returns twice as larger as those of a domestic bank. When the size effect dominates the difference in monitoring costs, the profit of the MNB dominates that of the domestic bank.

3.2. Complete bail-in ($\alpha = 0$)

We consider now the relative profitability of banks when public support is absent, that is, $\alpha = 0$. In this case investors require a return rate that fully incorporates the risk of the different business models. In particular, we have:

$$r_M^a(0) = \frac{1}{p(2-p)} < r_D(0) = \frac{1}{p} \tag{17}$$

Result in (17) follows from the diversification of the portfolio: bondholders in the MNB are repaid with a higher probability and thus accept a lower return rate. Now we can derive the expected profits of the different bank models.
In this case notice that all thresholds collapse to the same value:

$$c_D(0) = c_M(0) = c_\Pi(0) \tag{18}$$

Results in the case of $\alpha = 0$ can be summarized in the following proposition:

Proposition 2. *Assuming that each bank's monitoring cost is a specific realization from the uniform distribution on $[0, c_{Max}]$, that project return $R \geq \frac{2}{p(2-p)}$ occurs with probability of success $p \in (0.5, 1)$ and that returns are independent across countries, we have in the case of bail-in ($\alpha = 0$) that:*

(i) the minimum monitoring cost that guarantees positive expected profit for the two types of banks is the same;

(ii) when $c < c_D(0)$ the expected profit of a MNB is always higher than that of a domestic bank, implying that in the regions where the domestic bank is viable, the MNB is always more profitable.

Proof. Since the thresholds coincide for the two types of banks, both domestic and MNB banks are viable for the same set of realizations of c. The profit of the domestic bank is $E[\Pi_D(0)] = \max\left\{(pR-1) - \frac{c}{2}, 0\right\}$ while for a MNB is $E[\Pi_M^a(0)] = \max\left\{2(pR-1) - c, 0\right\}$. For $c \geq c_D(0)$, it is easy to prove that the profit of a MNB is larger than that of a domestic bank. □

Figure 3 shows the results in proposition 2. To understand the result in point (i) consider that for every unit of return, banks pay the same costs, whatever the structure they choose. This is so first because they pay different bond rates, but also since they repay creditors with different probabilities: with $\alpha = 0$, the riskiness of bonds is perfectly priced, therefore the two effects perfectly offset each others. Second, they face the same monitoring cost per

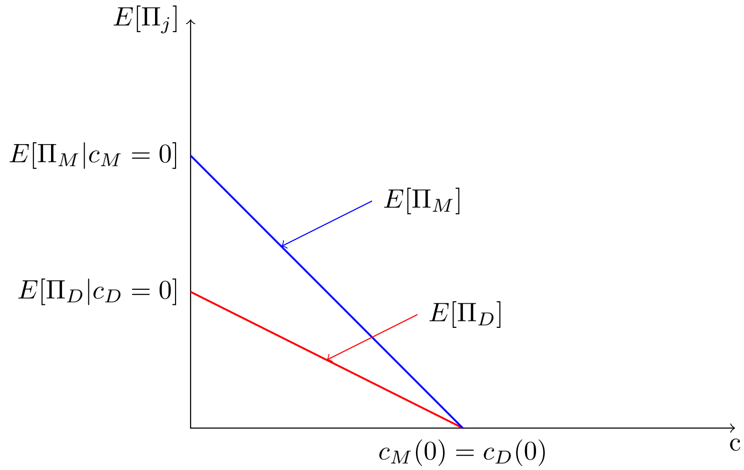

Fig. 3. Expected profits with $\alpha = 0$.

unit lent. The result at point (ii) follows from the result at point (i) since a MNB has the double the size of a domestic bank: whenever c is sufficiently low compared to the NPV of the project, i.e. $c < c_D(0)$, the expected profit of a MNB is the double compared to that of a domestic bank.

3.3. Comparative Statics

In this section we analyze the results obtained in sections 3.1 and 3.2 for different values of α. We recall here the expression of the different thresholds. First of all, the threshold on c for a domestic bank to be viable is:

$$c_D(\alpha) \equiv 2(pR - 1) + 2\alpha(1 - p) \tag{19}$$

while for a MNB:

$$\begin{cases} c_\Pi(\alpha) \equiv c_D(\alpha) - 4\alpha p(1 - p) \\ c_M(\alpha) \equiv c_D(\alpha) - 2\alpha p(1 - p) \end{cases} \tag{20}$$

We can illustrate the equilibrium outcome of the choice of banks in terms of the different realizations of c in Figure 3.

Figure 3 - here

The first effect relates to the change of the thresholds. In particular, for a generic level of $\alpha \in [0, 1]$ the impact on the thresholds of c for a MNB is:

$$\frac{\partial c_M(\alpha)}{\partial \alpha} = \frac{\partial c_D(\alpha)}{\partial \alpha} - 2p(1 - p) \quad < \quad \frac{\partial c_D(\alpha)}{\partial \alpha} = 2(1 - p) \tag{21}$$

From (21) we see that a rise in the level of α (moving from bail-in to bail-out) produces an increase in the thresholds, which is greater for the domestic bank than for the MNB. This is in line with the results obtained in the previous sections: with $\alpha = 1$ (bail-out) we have in proposition 1 that $c_M(1) < c_D(1)$, while with $\alpha = 0$ (bail-in) all the thresholds collapse to the same value. Indeed, with $\alpha = 1$ the bond rate is equal to 1 for both types of banks. However, the MNB bank repays bondholders with a higher probability, hence the funding cost per unit lent is higher compared to that of a domestic bank. As α decreases (from bail-out to bail-in), the bond rate increasingly prices the risk entailed by each type of bank. Due to the double direction in which cross subsidies can go in the case of a MNB, bondholders will be repaid with the highest probability and thus the bond rate for a MNB will increase relatively less. When $\alpha = 0$ the expected funding cost per unit lent is the same for all bank types since the bond rate fully reflects the different levels of risk. In this case the differences in the funding costs are exactly offset by the different probabilities with which the bond is repaid.

The second effect we need to consider is the effect on the profitability of banks. This effect is partly due to the different size of the two banks. In particular, on the one hand, the MNB earns the double of the revenues of a domestic bank. On the other hand, whenever $\alpha \neq 0$, the MNB pays a funding cost that is less than the double that of a domestic bank. Given the presence of public funds, the bond rates don't full reflect the risk entailed by the two business models. The domestic bank pays on average a lower bond rate, which is more than offset by the higher probability of repayment. In addition, the MNB faces higher operating costs. Overall the MNB earns higher expected profits when R is sufficiently high for the difference in revenues to overcome the difference in costs.

The following proposition generalizes the results in Propositions 1 and 2.

Proposition 3. *Assuming that each bank's monitoring cost is a specific realization drawn from the uniform distribution $[0, c_{Max}]$, and that project return $R \geq \frac{2}{p(2-p)}$ occurs with probability of success $p \in (0.5, 1)$ and that returns are independent across countries, we have that for any $\alpha \in [0, 1]$:*

- *(i) the MNB is more profitable than the domestic bank when $c < c_\Pi(\alpha)$;*
- *(ii) there exists a subset of realizations of c where the domestic bank is the most profitable, that is when $c_\Pi(\alpha) < c < c_D(\alpha)$ for any $\alpha > 0$;*
- *(iii) as the degree of public support α decreases, the MNB becomes more profitable relatively to domestic banks; in fact, as α decreases, the thresholds $c_\Pi(\alpha)$ decrease less compared to $c_D(\alpha)$: hence the mass of MNB increases compared to that of domestic banks.*

4. Optimal Resolution Policy

In this section we analyze the optimal resolution policy considering the outcome in terms of banks' structure. In particular, the analysis in section 3 shows that the structure of the banking sector is a function of the resolution policy announced by the regulator. Thus, in defining the optimal policy, the regulator will have to consider the implications on the structure of the banking system. At this point we need to focus on two effects:

- First, the level of α will influence the mass of viable banks, that is those that will not be able to raise funds and forced to close. This set of banks is defined by the threshold $c_D(\alpha)$.
- Second, it will influence the structure of banks that will find it profitable to switch from MNB to domestic, defined by the threshold $c_\Pi(\alpha)$.

To capture in reduced form the negative effect of banks' closure, we assume that the contribute of banks to the social welfare goes beyond its profits, and also includes the provision of social relevant services, i.e. payment services and the production of information on borrowers. This contribution is captured by the parameter $\gamma > 0$. We also assume that MNBs gives a double service, both in the domestic and the foreign country. This is to capture the fact that, when the regulator chooses a lower α (moving from a regime of bail-out to bail-in), the reduced provision of services caused by banks' closure can be substituted to some extent by the entry in the market of new MNBs.
Also, the reduction in the amount of loans caused by the closure of domestic banks, can be substituted by the increase in the amount of loans provided by MNBs, which are also better in picking profitable investment projects, given their ability to diversify risks.
However we assume that MNBs are more costly to supervise relatively to domestic banks due to the greater complexity of their activities across different countries: we therefore introduce a positive cost of supervisory complexity, that we define ψ per MNB bank.
The regulator maximizes the expected social welfare given by the sum of banks' profits, savings in taxpayers' money[3] plus the social value of the activity of each domestic bank less of the supervisory cost of complexity of MNBs:

$$E^D[W] = E[\Pi_D(\alpha)] + \gamma - \alpha(1-p) = pR - 1 - \tfrac{c}{2} + \gamma \tag{22}$$

whereas the contribution of each MNB to social welfare is:

$$E^M[W] = E[\Pi_M(\alpha)] + 2\gamma - 2\alpha(1-p)^2 - \psi = 2E^D[W] - \psi \tag{23}$$

[3]The model can be extended to include a social cost of distortionary taxation.

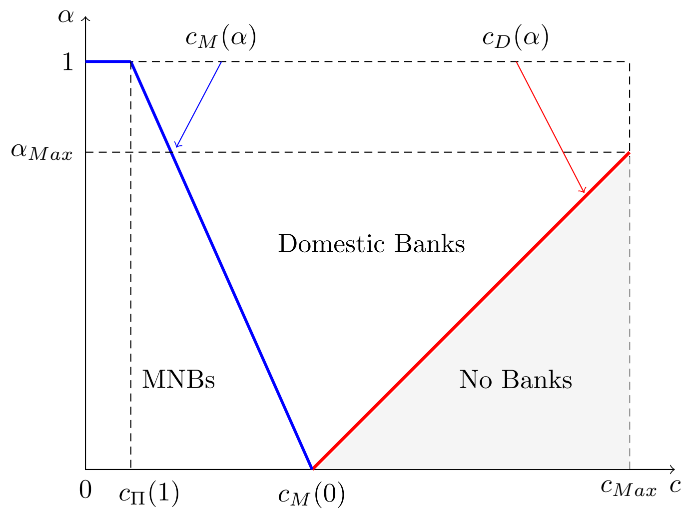

Fig. 4. Resolution policy: banks' business models and welfare.

The social contribution of MNBs is twice that of domestic banks due to the size effect, however there is an additional cost of supervision given by their complexity. Given this specification of the social welfare function, for a given banking structure, the resolution policy chosen by the regulator has no effects on the contribution of each bank to the social welfare, i.e. $\frac{\partial E^j[W]}{\partial \alpha} = 0$ for $j = \{D, M\}$ as can be seen from (22) and (23). However it changes the intervals of the regions where domestic and MNBs banks exist. In particular, MNBs will exist for c in the interval $[0, c_\Pi(\alpha)]$ while domestic banks in the interval $[c_\Pi(\alpha), c_D(\alpha)]$, where:

- $c_D(\alpha) = 2(pR - 1) + 2\alpha(1 - p)$;
- $c_\Pi(\alpha) = c_D(\alpha) - 4\alpha p(1 - p)$.

with $\frac{\partial c_D(\alpha)}{\partial \alpha} > 0$ and $\frac{\partial c_\Pi(\alpha)}{\partial \alpha} < 0$, implying that as the regulator increases α, it increases the fraction of domestic over MNBs. This implies that the effects on the structure of the banking system will be the only determinant of the level of social welfare. Indeed, we have that $|\frac{\partial c_D(\alpha)}{\partial \alpha}| < |\frac{\partial c_\Pi(\alpha)}{\partial \alpha}|$. Thus the mass of domestic banks closing down is larger compared to the new multinational banks entering the market. However the surplus generated by each MNB is the double compared to that of a domestic bank, although the cost of supervision of the MNBs is greater compared to domestic banks.

Figure 4 shows the relation between the resolution policy, the banking structure and the aggregate welfare.

The objective of the regulator is to choose the degree of bail-out that maximizes the social welfare:

$$E[W(\alpha)] = \frac{1}{c_{Max}} \int_0^{c_\Pi(\alpha)} \{2[pR - 1 + \gamma - \frac{c}{2}] - \psi\} \, \mathrm{d}c + \frac{1}{c_{Max}} \int_{c_\Pi(\alpha)}^{c_D(\alpha)} [pR - 1 + \gamma - \frac{c}{2}] \, \mathrm{d}c$$

Hence the regulator maximizes the following function w.r.t. α:

$$E[W(\alpha)]c_{Max} = (pR - 1 + \gamma)[c_D(\alpha) + c_\Pi(\alpha)] - \psi c_\Pi(\alpha) - \frac{1}{4}\left[c_D^2(\alpha) + c_\Pi^2(\alpha)\right] \tag{24}$$

deriving the optimal degree of bail-out:

$$\alpha^* = \frac{2(1-p)\gamma + (2p-1)\psi}{2(1-p)(1-2p+2p^2)} > 0 \tag{25}$$

In our simple framework the optimal resolution policy calls for a positive level of public support, given that both the numerator and the denominator in equation (25) are positive. The optimal degree of bail-out is increasing in the value of γ, that is the social value of financial services provided by banks and decreasing in the probability p, that is, how safe are banks. As the cost of supervisory complexity of MNBs increases, larger ψ, the greater is the optimal degree of bail-out tolerated in the economy.

5. Conclusions and Future Research

The analysis shows how banks' funding costs depend on the resolution policy chosen by the regulator. In particular, in line with recent empirical evidence, we show that "more" bail-in increases banks' cost of funding and its risk sensitivity. As a consequence, the relative expected profitability of different bank business models changes when we consider different resolution policies. The analysis has shown that, as we reduce the scope for public funds, there is an increased incentive for banks to expand abroad looking for new investment opportunities. Thus,as a result of the new resolution framework, we might see a stronger incentive for banks to internationalize, thus leading to grater financial integration. On the other hand, the increase in funding costs forces the least efficient banks to drop off the market. Building on the trade-off between the greater incentive to expand abroad, and the increased share of banks' closures, we identify the optimal resolution policy. In particular we show that in equilibrium it is optimal to have a positive level of public support, whose magnitude depends on the characteristics of the banking sector.

The results in our paper are consistent with the recent policy debate: [4] a diffuse opinion is that the new bail-in tool will have an impact on banks' funding costs and this might undermine the viability of smaller, less profitable banks. In this sense, the new resolution framework will improve the possibility of an orderly exit from the market of the least efficient banks, thus fostering consolidation in the European banking sector and potentially helping

[4] https://ftalphaville.ft.com/2019/01/16/1547653807000/How-much-will-it-cost-banks-to-borrow-/

the sector to increase its (subdued) profitability. [5].

The analysis can be extended in different ways: first, it would be interesting to introduce capital requirements. In this case, we can see the interaction between the role played by *ex-ante* capital regulation and *ex-post* resolution policies. In addition, the introduction of a risk-weighted capital requirement could further expand the scope for diversification. Another direction in which the analysis can be extended is to add heterogeneity among regulators in terms of attitude towards public support to see the implications for banks' incentives to expand abroad and for the possible creation of a supranational banking union.

[5] https://www.bankingsupervision.europa.eu/press/speeches/date/2019/html/ssm.sp190704 1f442782ac.en.pdf

References

Acharya, V., Drechsler, I., and Schnabl, P. (2014). A pyrrhic victory? Bank bailouts and sovereign credit risk. The Journal of Finance, 69(6), 2689-2739.

Aldasoro, I., Hardy, B., and Jager, M. (2020). The Janus Face of bank geographic complexity. *Working Papers*, 858, BIS.

Avgouleas, E., and Goodhart, C. (2015). Critical reflections on bank bail-ins. Journal of Financial Regulation, 1(1), 3-29.

Crespi, F., Giacomini, E., and Mascia, D. V. (2019). Bail-in rules and the pricing of Italian bank bonds. European Financial Management, 25(5), 1321-1347.

Cutura, J. (2018). Debt holder monitoring and implicit guarantees: Did the BRRD improve market discipline? *Working Paper*.

Dell'Ariccia, M. G., Peria, M. M. S. M., Igan, M. D. O., Awadzi, E. A., Dobler, M., and Sandri, M. D. (2018). Trade-offs in Bank Resolution. International Monetary Fund.

Deng, S., and Elyasiani, E. (2008). Geographic diversification, bank holding company value, and risk. Journal of Money, Credit and Banking, 40(6), 1217-1238.

Dewatripont, M. (2014). European banking: Bailout, bail-in and state aid control. International Journal of Industrial Organization, 34, 37-43.

Duijm, P., and Schoenmaker, D. (2020). European banks straddling borders: risky or rewarding?. Finance Research Letters, 101525.

Faia, E., Ottaviano, G. I., and Sanchez, I. (2017). International expansion and riskiness of bank, *Working Paper*.

Fang, Y., and van Lelyveld, I. (2014). Geographic diversification in banking. Journal of Financial Stability, 15, 172-181.

Fiordelisi, F., and Scardozzi, G. Bank Funding Strategy After the Bail-in Introduction, *Working Paper.*

Giuliana, R. (2019). Impact of Bail-In on Banks' Bond Yields and Market Discipline, *Working Paper.*

Goetz, M. R., Laeven, L., and Levine, R. (2016). Does the geographic expansion of banks reduce risk?. Journal of Financial Economics, 120(2), 346-362.

Laeven, L., and Valencia, F. (2013). Systemic banking crises database. IMF Economic Review, 61(2), 225-270.

Levine, R., Lin, C., and Xie, W. (2016). Geographic diversification and banks? funding costs, *Working Paper*, w22544, National Bureau of Economic Research.

Lewrick, U., Serena, J. M., and Turner, G. (2019). Believing in bail-in? Market discipline and the pricing of bail-in bonds, *Working Paper.*

Lintner, P., Lincoln Nagy, M. A. J., Pyziak, P., Godwin, A. J., Schroeder, S. C., and Irsalieva, N. (2016). Understanding bank recovery and resolution in the EU: a guidebook to the BRRD (No. 112266, pp. 1-182). The World Bank.

Luciano, E., and Wihlborg, C. (2018). Financial synergies and systemic risk in the organization of bank affiliates. Journal of Banking and Finance, 88, 208-224.

Pandolfi, L. (2018). Bail-in vs. Bailout: a False Dilemma?, *Working Paper*, 499, Centre for Studies in Economics and Finance (CSEF), University of Naples.

Philippon, T., and Salord, A. (2017). Bail-ins and bank resolution in Europe. Geneva Reports on the World Economy Special Report, 4.

Schfer, A., Schnabel, I., and Weder, B. (2016). Bail-in expectations for European banks: Actions speak louder than words, *Working Paper.*

Walther, A., and White, L. (2019). Rules versus discretion in bank resolution. The Review of Financial Studies.

Appendix A. Subsidiaries

When expanding abroad, banks can choose to do so via a foreign subsidiary. Differently from the branch structure, in this case the subsidiary is responsible for the parent's liabilities while the converse is not true. This implies that the probability of being repaid is different for domestic or foreign depositors. As such, the deposit rate that satisfies their participation constraint is different. In particular, home depositors can expect to be repaid both when only the home project is successful and when only the foreign one succeeds. On the contrary, foreign depositors can expect to be repaid only when the foreign project is successful. Consequently we define r_H the deposit rate offered to depositors in the home country and r_F the deposit rate offered to depositors of the foreign unity. Taking this into account, the expected profits of a multinational bank that operates with a subsidiary can be written as:

$$E[\Pi_M^S] = p^2\ max\{2R - r_D - r_F, 0\} + p(1-p)\ max\{R - r_D\} + p(1-p)\ max\{R - r_F - r_D\} - c \quad (26)$$

The second term in the profit refers to the case where only the domestic project is successful and the other one fails (with probability $p(1-p)$): in this case the revenues from the project are not used to repay foreign depositors. Conversely, as shown by the third term, when the foreign project is successful, the revenues are used to repay both domestic and foreign depositors. Thus, in this last case, we have to distinguish between two cases:

- case (a), when $r_D + r_F \leq R$: the return from the successful project is enough to repay the promised rate to bondholders;
- case (b), when $r_D + r_F > R$: what is returned by the successful project is not enough to refund all bondholders; in this case $(1-\alpha)$ bondholders receive $\frac{R}{2}$, while another fraction α is bailed-out and receive 1 unit as public support.

In line with the case of branch MNB, we introduce the following assumption:

$$R \geq \frac{3-p}{p(2-p)} \quad \text{(A3)}$$

With A3,we assume that the revenues from the successful projects are always high enough to repay both domestic and foreign depositors, independently from the resolution policy chosen by the regulator . In this case, the two participation constraints are:

$$\begin{cases} p(2-p)r_H + (1-p)^2\alpha = 1 & for\ domestic\ investors \\ pr_F + (1-p)\alpha = 1 & for\ foreign\ investors \end{cases} \quad (27)$$

This implies that the required return rates are:

$$\begin{cases} r_H = \frac{1-\alpha}{p(2-p)} + \alpha \\ r_F = \frac{1-\alpha}{p} + \alpha \end{cases} \quad (28)$$

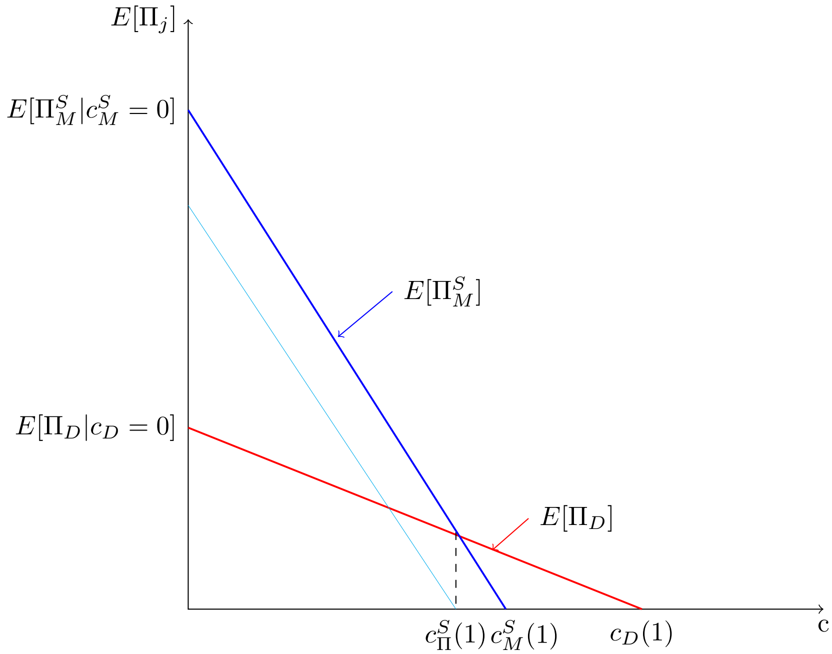

Fig. 5. Expected profits with $\alpha = 1$, including subsidiaries.

Plugging the return rates defined in (28) into (26) we obtain the expected profits of the subsidiary MNB:

$$E[\Pi_M^S] = 2pR - 2 + 2\alpha - 3p\alpha + p^2\alpha - c \tag{29}$$

We can now derive the threshold of c below which a subsidiary MNB is profitable:

$$c \leq c_M^S(\alpha) \equiv c_D(\alpha) - p\alpha(1-p) \tag{30}$$

From (30) we can notice that $c_D(\alpha) > c_M^S$: similarly to the case of a branch MNB, the subsidiary MNB needs to be more efficient than the domestic bank in order to be viable. This is again explained by the fact that the domestic bank maximizes the expected value of public money, while the (partial) effect of diversification in the subsidiary MNB reduces its value. Furthermore, we can notice that $c_M^S > c_M^B$: similarly to what observed above, we can now say that the partial cross-liability embedded in subsidiaries increases the expected value of public subsidies compared to the full cross-liability envisaged for branches. As such, a branch MNB needs to be more efficient than a subsidiary MNB in order to be viable.

Finally, we can compare the profits of subsidiary MNBs with those of domestic banks. This will allow to define the interval of values of c where subsidiary MNBs are the most profitable bank structure. Comparing equation (29) with (5), we obtain:

$$c_\Pi^S(\alpha) \equiv c_D(\alpha) - 2\alpha p(1-p) \tag{31}$$

From relation (31) we can see that, similarly to the case of branch MNBs, whenever $\alpha > 0$, the threshold of c below which subsidiary MNBs are the most profitable bank structure is lower that $c_D(\alpha)$. As such, if we move a way from a situation of full bail-in, there will be an interval of realizations of c, i.e. $c_\Pi^s < c < c_D$, where domestic banks are the most profitable.

Figure 5 compares the profitability of subsidiary MNBs and domestic banks with $\alpha = 1$. On the contrary, if the regulator chooses a full bail-in policy, subsidiary MNBs always dominate domestic banks, whenever the two banks are viable. In this sense, figure 3 is suitable to compare the profitability of subsidiary MNBs and domestic banks.

Fire Sales, Balance Sheet Degradation, and the Signaling Role of Illiquidity

Davide Bosco* Stefano Montoli †

November 2020

Abstract

Insolvency risk and (il)liquidity risk are key determinants of the fragility of financial institutions. This paper draws a connection between the two: through a stylized model, we show how an a priori sound institution can be pushed into insolvency *because of* the presence of an informationally efficient market for financial assets. Due to temporary liquidity needs, a financial institution starts selling some of its assets, signalling to the market that its fundamentals may be bad. The consequent reduction in buyers' willingness to pay deflates asset prices, thus forcing the institution into a fire sales spiral, that might push the institution into insolvency. The nature of the investors' strategic interaction is determined endogenously: strategic complementarity arises when investments' cash flow realizations, determined by the fundamentals and temporary shocks, are sufficiently bad. Conversely, substitutability prevails when the generated cash flows are sufficiently strong.

JEL classification: D82, D84, G23, G28.

Keywords: Coordination, Feedback effect, Financial-market runs, Heterogeneous information, Strategic complementarities.

*Catholic University of Milan. Email: davide.bosco@unicatt.it

†Bicocca University and Center for European Studies (CefES). Email: s.montoli3@campus.unimib.it

1. Introduction

The sudden burst of the subprime mortgage bubble in 2007, and the subsequent turmoil in credit markets induced by the rapid spread of panic among major market agents, highlighted how a better understanding of the mechanics at work beneath financial fragility crucially depends on the ability to capture the peculiar balance sheet composition of financial intermediaries. As the unanticipated shock of the subprime crisis undermined investors' confidence, the evaporation of short-term funding induced by widespread panic-based runs rapidly turned liquidity issues into fundamental distress. Peculiar in this respect is the run on the Special Investment Vehicles (SIVs) of 2007 cited by Vives (2014). Perceiving an increased counter-party risk due to the high exposure of SIVs to the adverse effects of write-downs on mortgage-related securities, money-market funds started to refuse rolling-over their short-term funding. The liquidity shock to SIVs conduits quickly turned into distress for sponsor commercial banks (see Brunnermeier (2009) for a detailed account of the issue). Among the several lessons to be learned from the panic of 2007, and the related liquidity crunch of 2007-2008, one particularly relevant to the analysis presented in this paper is that bad fundamentals may result from illiquidity[1]. While great effort has been dedicated by the traditional literature on bank runs to the analysis of the mechanisms that turn bad fundamentals into illiquity, relatively less attention has been devoted to illiquidity-based insolvency. Among the others, Brunnermeier and Pedersen (2009), Vives (2014), Morris and Shin (2016) and Matta and Perotti (2018) are notable exceptions. Morris and Shin (2016), in particular, call for a joint analysis of illiquidity and insolvency risk.

In the spirit of Morris and Shin (2016), we provide a tractable model where liquidity and insolvency risk are naturally intertwined and jointly analyzed. In particular, illiquidity feeds into insolvency because of fire-sale spirals induced by the former. We analyze a dynamic market game where fire sales of assets, induced by temporary liquidity needs, generate a deflationary pressure on asset market prices, whereby a balance-sheet degradation cycle arises (see Figure 1). As a leading example, we focus on a stylized environment where the shareholders of an open-ended investment fund are called to decide whether to ask for immediate redemption of their holdings, or to wait until the liquidation of the fund at a subsequent date, after having *privately* observed information about the current fund's performance. Upon receiving investors' requests for redemptions, the fund manager sells a corresponding quantity

[1] Vives (2014) effectively highlights the issue as follows:

> "[...] *liquidity evaporates when short-term investors rush to exit, after which time a solvency problem may arise.*"

Quoted from Vives, X. (2014), "Strategic Complementarity, Fragility, and Regulation". *Review of Financial Studies*, 27, pag. 3547.

of the assets on the market. Investors bear the price risk under both courses of action, as their net payoff coincides with the prevailing market value of their holdings – i.e. the Net Asset Value (NAV) of the fund, divided by the number of outstanding shares. Since early redemptions are more likely when the fund's performance is poor, a potential purchaser become skeptical upon observing a large supply of assets sold. Lower willingness to purchase by market agents drives asset prices down, in turn signaling to the whole market the fund's poor performance. As a result the fund's liquidation value drops. Since both current and future liquidation price are negatively affected by early redemptions, it is ex ante unclear whether individual actions are complements or substitutes. We prove that complementarity dominates when the cash flows generated by the investment portfolio are sufficiently low. When early redemptions by individual investors are strategic complements, early redemptions by pessimistic[2] investors drain resources available to patient ones and, at the same time, erode their market value, thus precipitating a deflationary spiral that induces other, less pessimistic investors to redeem their share, eventually pushing the fund into insolvency. When substitutability prevails, bad news may induce early redemptions, but do *not* generate a balance-sheet degradation cycle. Both anecdotal and empirical evidence seem to favor the complementarity hypothesis. Liu and Mello (2011) quote the following excerpt from a 2009 report on hedge funds by the IFSL[3]:

> "*Hedge funds faced unprecedented pressure for redemptions in the latter part of 2008, with investors withdrawing funds due to dissatisfaction with the performance or to cover for even greater losses or cash calls elsewhere. This in turn led to forced selling and closures of positions by hedge funds causing a cycle of further losses and redemptions.*"

Analyzing data on mutual funds in the years 1995-2005 from the CRSP[4] database, Chen et al. (2010) provide sound empirical evidence that points to the presence of strategic complementarity among individual redemption decisions. Employing data from the same database, Coval and Stafford (2007) prove that mutual funds' asset sales in correspondence of large capital outflows (redemptions) exert significant pressure on prices of sold securities. Our choice to focus on coordination failures that potentially arise in the realm of open-ended investment schemes, rather than 'traditional' runs on commercial banks, is driven by three main motivations. First, as extensively discussed in Rochet and Vives (2004) and

[2]Throughout the paper we define pessimistic investors as the individuals whose private information is sufficiently bad that they decide to liquidate their shares as soon as they have the opportunity to do so. As such, the label pessimistic refers to their course of action throughout the game. On the contrary, we don't refer to any behavioural characteristics which might translate in specific features of their utility function.

[3]International Financial Services London Research.

[4]Center for research in security prices.

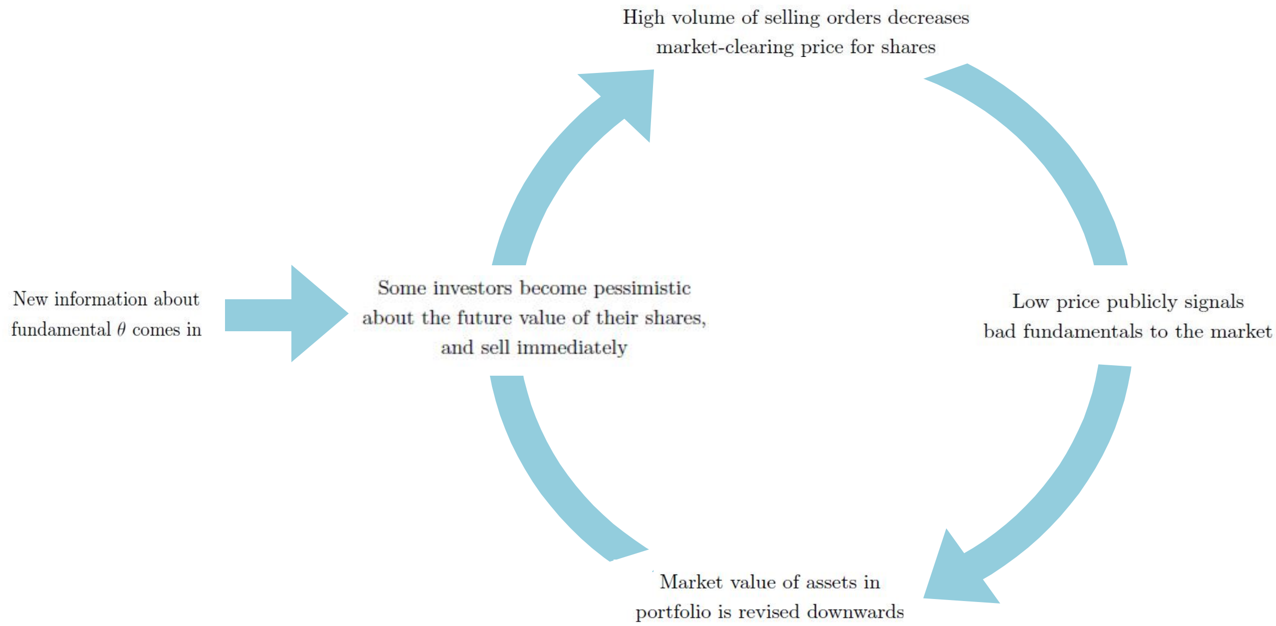

Fig. 1. Diagrammatic representation of a balance-sheet degradation cycle induced by fire sales. The informational externality implicitly generated by the market-clearing price of shares – sold by more pessimistic primary investors on the secondary market – directly affects the mark-to-market valuation of the assets in portfolio, in turn fostering additional fire sales from less pessimistic primary investors.

Vives (2014), when 'modern' bank runs are to be analyzed, the explanatory power of the deposit-based framework of traditional bank runs á la Bryant (1980) and Diamond and Dybvig (1983) may be somehow limited. Lender-of-last-resort facilities (LOLR)[5] and deposit insurance schemes (DISs), widely adopted by developed economies after the financial crisis of 1929, proved successful in containing deposit-based bank panics[6]. Moreover, as highlighted by Brunnermeier (2009), during the years that preceded the great financial crisis of 2007/2008, the banking industry witnessed a general tendency towards an 'originate-and-distribute' business model, coupled with a sharp increase in the usage of short(er) maturity instruments to finance asset holdings. As a consequence of these (recent) trends in the banking industry, overall systemic fragility increasingly depends on the fragility of institutional liquidity providers such as money-market funds. Factual evidence in this respect is not hard to find in the recent past. The adverse – and long-lasting, see Kabir and Hassan (2005) – systemic effects of the (quasi) collapse of Long-Term Capital Management (LTCM)

[5]Fed interventions after the collapse of Penn Central in 1970 and after the stock market crash of 1987 can be considered as recent examples of LOLR facilities.

[6]See Demirgüç-Kunt et al. (2015) for a comprehensive account of DISs as of 2013.

in 1998 prompted an intervention of NY Fed, that openly sponsored a concerted bail-out of about $ 3.625 billion by the major creditors[7]. More recently, the 2007 subprime crisis fostered a run on Asset-Backed Commercial Paper (ABCP) conduits and Special Investment Vehicles (SIVs), that significantly contributed to the onset of the financial turmoil generally referred to as the '2007 panic' (see e.g. Gorton and Metrick (2012)). Notably, the refusal of (many) money-market funds to roll-over their short-term lending to ABCP programs played a key role in the spread of the panic[8]. In 2008 hedge-funds, too, had to face an unprecedented volume of redemption by their investors. Second, in light of the increased reliance of both commercial and investment banks on money-market wholesale funding, a better understanding of the behavior non bank financial institutions is of particular interest – especially as the deep mechanics that governs money markets substantially differ from other – e.g. stock – markets (see Holmström (2015) for a detailed discussion of the issue). To the best of our knowledge, only Liu and Mello (2011) and Vives (2014) explicitly analyze a coordination problem among the investors of a non-bank financial intermediary. It is therefore important to improve our understanding of how (some of) the peculiar organizational characteristics, that distinguish non-bank intermediaries from traditional commercial banks affect their resilience to panic-induced runs by investors. Third, consideration of an institutional arrangement where capital providers bear price risk simplifies both the analysis and the derivation of results in a non-trivial manner, while retaining an high comparability with standard (commercial) bank-run models.

Furthermore, our choice to focus on open-ended financial intermediaries is firstly determined by the consideration that they account for the vast majority of modern collective investment vehicles – see e.g. Stein (2005). Mutual funds, money-market funds and hedge funds, are all examples of open-ended investment schemes[9]. In addition, open-ended financial institutions are more prone to self-fulfilling, panic-induced runs than their closed-ended counterparts, for their capital structure is, on average, more fragile. Indeed, albeit fund managers of open-ended schemes can often impose restrictions on the (early) redemption of shares, investors are often allowed to liquidate their positions on (relatively) short notice – see Giannetti and Kahraman (2017) and Dimson and Kozerski (2002)[10].

[7]See Edwards (1999) and Lee (2018) for a more detailed account of the issue.

[8]ABCP are structured products that result from the securitization of commercial banks' mortgages. After the burst of the subprime bubble, skepticism spread among liquidity providers of ABCP conduits about the ability of the latter to repay short-term debt. As a consequence, massive amount of debt could not be rolled-over. Notably, money-market funds were the most important providers of short-term liquidity to ABCP conduits.

[9]European SICAVs (Société d'Investissement à Capital Variable) and UTs (Unit Trusts) are other examples of open-end investment vehicles.

[10]Dimson and Kozerski (2002), in particular, provide a comprehensive comparative analysis of the main

Structure of the paper. The present work is structured as follows. In Section 2 we review (some of) the theoretical literature pertinent to our analysis. In Section 3 we present the general model and extensively discuss its primitives. In Section 4 we provide a fully analytical characterization of the unique monotone equilibrium of the model. In Section 5 we derive and discuss some additional results.

2. Literature Review

Our work relates to and contributes four main strands of theoretical literature. First, we contribute to the classical literature on bank runs, pioneered by Bryant (1980), Diamond and Dybvig (1983) and Postlewaite and Vives (1987). Liu and Mello (2015) highlight that several similarities exist between the demand-deposit contracts offered by commercial banks, and the equity capital issued by (most) non-bank, open-end institutions – for the latter can often be redeemed by investors on short notice[11]. Hence, albeit designed for non-bank financial institutions, our framework of analysis – as well as *some* of our main results – readily extends to more standard, bank-like financial intermediaries. In the spirit of Morris and Shin (2000, 2003 and 2004), Goldstein and Pauzner (2005) and Allen et al. (2018), we build on the global games approach of Carlsson and van Damme (1993), and study individual investment decisions within a coordination game of incomplete information, where decentralized decision-making is driven by self-fulfilling, rational beliefs, and where beliefs are, in turn, pinned down by underlying (unobserved) economic fundamentals. Similarly to the aforementioned contributions, the uniqueness of the monotone equilibrium of our model hinges on incomplete (asymmetric) information. Differently from these contributions, however, we do *not* impose strategic complementarity *a priori*. In our model, strategic complementarity arises *endogenously* as a consequence of the (*ex ante*) optimal selling/purchasing decision of rational, incompletely informed investors. Within the vast literature on bank runs, closest in spirit to our work are Rochet and Vives (2004), Vives (2014), Eisenbach et al. (2014) and

characteristics of open- and closed-ended structures.

[11] The authors underline that (page 492)

> *"One key feature of the capital structure of hedge funds is the fragile nature of their equity. Equity capital in hedge funds can be redeemed at investors' discretion, a feature somewhat similar to demand deposit-dept in banks".*

Recall that the regulatory framework of hedge funds differs substantially from that of other open-ended financial institutions. Differently from mutual fund, hedge funds are allowed to restrict investors' redemptions via contractual provisions such as redemption-notice periods and payout periods (see Giannetti and Kahraman (2017)). Hence, the fragility argument outlined by Liu and Mello (2011) with respect to hedge funds extends, *a fortiori*, to mutual and money-market funds. For a detailed comparison between closed-ended and open-ended organizational structures, see Dimson and Kozerski (2002).

Morris and Shin (2016). The authors jointly analyze illiquidity and insolvency by focusing on the peculiar balance sheet composition of financial intermediaries. Rochet and Vives (2004) and Vives (2014), in particular, outline models where temporary liquidity needs force financial intermediaries to sell (some of) their assets at fire-sale prices. The consequent, self-sustaining balance-sheet degradation cycle, in turn, potentially triggers their insolvency – Liu and Mello (2015) employ a similar modeling approach to study (the fragility of) hedge funds. Similarly to these contributions, we directly link (il)liquidity to (in)solvency via fire-sales. Differently from those contributions, however, we consider a model where the fire-sale discount on the fundamental value of assets is determined *endogenously* by the signaling effect embedded into early-selling decisions[12].

A key element of our model is the presence of a financial market that, albeit competitive and informationally efficient, allows for the emergence of fire-sale discounts, hence for the (temporary) mispricing of assets. In this respect, our paper relates to the literature on limited arbitrage that builds on the seminal work of Shleifer and Vishny (1997). Differently from most contributions in this strand of literature, we do *not* consider fire-sale prices as inefficient byproducts of the unwillingness of major arbitrageurs to bet against market underpricing. Instead, we interpret fire-sale discounts as pecuniary externalities, that arise from coordination failures in the decentralized decision-making of rational, incompletely informed investors[13]. In this respect, closest in spirit to our work are Brunnermeier and Pedersen (2009) and Matta and Perotti (2017). Matta and Perotti (2017) consider a model of bank run where temporary liquidity needs force a financial intermediary to sell some of the assets in its portfolio, and where the liquidation price is directly affected by the *exogenous* (unobserved) liquidity of the market as a whole. By abstracting from asset-specific components, the authors implicitly assume that the prevailing market conditions entirely determine the instantaneous liquidity of the asset sold. As a consequence, investors' aggregate selling and purchasing decisions affect the liquidation price via market-clearing, but not the intrinsic liquidity of the asset. Similarly to Matta and Perotti (2017), we study a bank run model where the liquidation price of an asset – that determines the individual incentives to run – is determined by its liquidity. However, in the spirit of Brunnermeier and Pedersen (2009), we model the overall liquidity of the asset as the sum of two, distinct components: a (i) market-wide and an (ii) asset-specific liquidity component. The former is purely exogenous – as in Matta and Perotti (2017) –, the latter is *endogenous*. The exogenous, market-wide liquidity component parameterizes the aggregate market demand for the asset. The asset-

[12]Highly realistic balance sheet structures are analyzed by Eisenbach et al. (2014). The relative complexity of their setup however precludes the use of a global-game refinement, so that the model exhibits multiple equilibria.

[13]See Shleifer and Vishny (2011) for an excellent review of the literature on fire sales.

specific component of liquidity is instead determined by the endogenous selling decisions of investors, that convey to other market participants (public) information about the unobserved fundamentals of the asset sold. Potential buyers extract information about the fundamentals of the asset from the observation of the aggregate market supply, and update their individual evaluations accordingly. As a consequence, under- and/or over-pricing arise endogenously.

Third, we contribute to the literature on signaling and endogenous learning in global games. Within this vast strand of literature, closest in spirit to our work are Hellwig et al. (2006), Dasgupta (2007), Tarashev (2007), and Goldstein et al. (2011, 2013). The authors study dynamic coordination games, where sophisticated players that take action in later stages of the game extract information upon the – private or public – observation of co-players' behavior in previous stages. Aware of the signaling effect exerted by their *aggregate* behavior, early movers internalize the informational externalities and choose accordingly their optimal courses of action. In particular, the information structure considered in our model heavily builds on that envisaged by Dasgupta (2007) in his sequential investment game, where late movers receive exogenous, noisy information about a non-linear transformation of (unobserved) early movers' aggregate action. Furthermore, similarly to Goldstein et al. (2013) and Hellwig et al. (2006), the aggregate market supply is specified in our model as a cumulative density function (CDF). However, while in Goldstein et al. (2013) and Hellwig et al. (2006) such an exotic functional form is adopted as a primitive of the model – and is essentially devised in order to improve analytical tractability –, in our model the CDF specification arises *endogenously*, as a direct consequence of buyers' individual optimal choices. As in Goldstein et al. (2013), the Walrasian auction that clears the market is efficient at aggregating traders' private information so that, in the spirit of Hayek (1945), the resulting clearing price constitutes *de facto* an unbiased, noisy signal, informative about the underlying, unobserved fundamentals of the asset(s) traded in the market. Therefore, as in Angeletos and Werning (2006), Tarashev (2006), and Amador and Weill (2010), late movers engage in *learning from prices*. However, since in our model late movers' price-contingent behavior directly affects initial movers' (expected) utilities, within a relatively simple framework we are able to analyze a self-sustaining negative feedback loop, where bad information signaled to the market by early redemptions corrodes the net worth of remaining shareholders, thus fostering panic and further redemptions.

Finally, our model indirectly builds on the demand-rotation framework, sketched by Lewis and Sappington (1994) in an early application to monopolistic price discrimination, and subsequently fully developed by Johnson and Myatt (2006). Similarly to these contributions, our model features an aggregate demand function whose endogenous price elasticity is determined by the underlying dispersion of (potential) purchasers' individual willingness to pay, and is therefore sensitive to the precision of available information. Cumulative density functions are then easily interpreted as inverse demand function. While our model does not involve any direct rotation of the aggregate market demand, it is possible to show that such an interesting feature can be obtained via a minor manipulation of the information structure.

3. The Model

3.1. General Setup

We consider a stylized financial system, populated by a continuum of atomistic, rational agents of total mass $m \gg 2$. Two main institutions operate within the system: (i) a financial institution, structured as an open-end collective investment scheme – henceforth, the *(investment) fund* –, and (ii) a market for financial assets, where securities are traded without any form of intermediation. The whole system lasts indefinitely in time, but the fund operates only over three periods, indexed by $t \in \{0, 1, 2\}$. Throughout the paper, we refer to $t = 0$ as the *initial date*, to $t = 1$ as the *interim date*, and to $t = 2$ as the *final date*. The financial market is assumed to be (i) competitive and (i) informationally efficient. Both characteristics of the market are *common knowledge*. The first assumption entails that no single trader can affect market prices, and stems directly from the atomistic nature of the economic agents. A direct consequence of the common knowledge of market competitiveness is that it is rational for all traders to act as pure price-takers. The second assumption entails that, besides guaranteeing the efficient allocation of financial assets, market prices aggregate and publicly disclose – at least in part – the private information possessed by individual traders.

The fund collects financial resources by issuing new *shares* via a primary emission, and (re)invests them into a portfolio of risky assets, that generate stochastic cash flows over an infinite time horizon. In reason of its non-bank nature, the fund is allowed to collect equity capital but *not* to lend financial resources to agents (e.g. by extending loans). All the newly-issued shares are offered to (a subset of) the atomistic agents and, once underwritten, can be (re)sold into a *niche* of the financial market at a subsequent date. Henceforth, we refer to the niche as the *secondary market* for fund's shares. Each atomistic agent is randomly assigned to one of three disjoint clusters: (i) *primary* investors, (ii) *secondary* investors, and (iii) residual market agents. Primary investors are of total mass $f = 1$, uniformly distributed over the unit interval $[0, 1]$ and indexed by i. They constitute, at the same time, the demand side of the primary market[14] and the supply side of the secondary market. At the initial date $t = 0$, every individual primary investor can either participate or not into the fund's (primary) emission. Upon participation, she is allowed to underwrite at most one single share. At the subsequent (interim) date $t = 1$, every underwritten share can be redeemed at its instantaneous market price by placing an *early-selling order* to the fund manager. Regulatory restrictions on early redemptions mandate that individual early-selling orders (i) must be submitted to the fund manager *in advance*, and (ii) cannot be in price-contingent form. As a consequence, primary investors do *not* know the (instantaneous) liquidation price of their shares at the moment they place their selling orders. Therefore, early redemption of shares is intrinsically *risky*. Secondary investors are of total mass $s \geq 1$, uniformly distributed over $[0, s]$ and indexed by j. At the interim date $t = 1$, they can purchase (or not) the shares redeemed by primary investors and sold by the fund manager in the secondary market. As a consequence, differently from primary investors, they are allowed to place price-contingent (purchasing) orders. Finally, residual agents are of mass $r \gg 2$, uniformly distributed over $[0, r]$, and indexed by ℓ. At the final date $t = 2$ they observe the market-clearing price of the secondary market, but they never actively participate in trade in that market, at any date $t \in \{0, 1, 2\}$. Clusters are unambiguously identified by the index variable $\iota \in \{F, S, R\}$, where (i) F indicates the cluster of primary investors, (ii) S indicates the cluster of secondary investors, and (iii) R indicates the cluster of residual market agents.

Individual types are observable, hence *common knowledge*, and agents cannot move across clusters. The total mass m of the atomistic agents in the economy can finally be written as

$$m = 1 + s + r \ ,$$

with $1 + s \geq 2$ by construction[15]. Note that, by assuming that $r \gg 2$, we are essentially imposing that the mass of transactions in the secondary market for the fund's shares accounts for a sufficiently small portion of the total volume of transactions in the financial market as a whole. All atomistic agents in the economy are *risk neutral*. For the sake of simplicity, we assume that they derive utility directly from the possession of financial resources[16]. At every date $t \in \{0, 1, 2\}$, they can lend (resp. borrow) financial resources by purchasing (resp. issuing) a safe asset, that yields the risk-free interest rate $\delta > 0$ and has a natural maturity of one period[17]. In particular, we assume the existence of a funding market for investors, which

[14]The investment fund implicitly defines the supply side of this reduced-form primary market.

[15]Recall indeed that $s \geq 1$ by assumption.

[16]Alternative, we can assume that the atomistic agents derive utility from the consumption of (instantaneous) quantities of a single commodity, whose price is constant over the their three-period lifetime and normalized to one.

[17]In other words, if an agent borrows one unit of money in t, she must pay an amount $1 + \delta$ at the

accommodates all investors' financial needs. We further assume that investors cannot engage in direct transactions among themselves, i.e. to lend or borrow resources they need to turn to the funding market. All financial transactions that involve the direct participation of the agent entail a *per-unit* (sunk) implementation cost $c > 0$, assumed positive but arbitrarily small[18]. The instantaneous financial endowment $w_t^{\mathbb{H}(\iota)} \in \mathbb{R}$ possessed by the generic agent $\mathbb{H}(\iota)$ at date t is mapped into utility via a monotone function $u : \mathbb{R} \longmapsto \mathbb{R}$, defined as

$$u\left(w_t^{\mathbb{H}(\iota)}\right) = w_t^{\mathbb{H}(\iota)} , \tag{1}$$

common to all agents and time-invariant, and with

$$w_0^{\mathbb{H}(\iota)} = 0 , \tag{2}$$

for every agent $\mathbb{H}(\iota)$ – i.e. agents do not possess financial resources at the initial date $t = 0$[19]. All atomistic agents are rational, hence they choose their courses of action in order to maximize the discounted sum of all future (expected) financial endowments (over their three-period lifetime). We call *total (expected) utility* the discounted sum of all future financial endowments. The individual utility is assumed *separable in time*, and time preferences are summarized by a discount factor $\Lambda \in (0, 1)$, common to all agents. At any arbitrary date $T < 2$, the total expected utility $EU_T^{\mathbb{H}(\iota)}$ of the generic agent $\mathbb{H}(\iota)$ can be defined as

$$EU_T^{\mathbb{H}(\iota)} = \mathbb{E}\left[\sum_{t=T}^{2} \Lambda^{t-T} w_t^{\mathbb{H}(\iota)} \,\middle|\, \mathcal{I}_T^{\mathbb{H}(\iota)}\right] , \tag{3}$$

for every $T \in \{0, 1\}$ and every cluster $\iota \in \{F, S, R\}$, where $\mathcal{I}_T^{\mathbb{H}(\iota)}$ is the information set of the generic agent $\mathbb{H}(\iota)$ at the arbitrary date T. In words, the total expected utility of the generic agent $\mathbb{H}(\iota)$ over her entire three-period lifetime, evaluated at the arbitrary date $T = 0$ in light of all information available, coincides with the discounted sum of all her future (expected) instantaneous amounts of consumption $q_t^{\mathbb{H}(\iota)}$ over the remaining periods

subsequent date $t + 1$. If an agent lends one unit of money in t, she receives *for sure* an amount $1 + \delta$ at the subsequent date $t + 1$.

[18]The small implementation cost c can be interpreted as an implicit cost of effort, imposed onto the agent by the necessity to arrange and manage the transaction. Since c is a transaction cost *per unit*, an amount $1 - c$ of financial resources is effectively invested for every unit of (gross) nominal investment.

[19]We introduce this assumption in order to improve the tractability of the model, as it simplifies the comparison of players' payoffs for at later stages of the game. Please notice that this assumption does not entail any loss of generality for the results presented in the paper.

$t = T, \ldots, 2$. For consistency, we finally impose the following *no-arbitrage condition*

$$\Lambda = \frac{1}{1+\delta} \ , \tag{4}$$

that guarantees that, at any date $t \in \{0, 1, 2\}$, no agent can realize a positive net profit by borrowing financial resources at the risk-free interest rate, and (re)investing them into the safe asset. Indeed, recall from expression (2) that agents have no initial endowments of financial resources, hence they *must* borrow money – at the risk-free interest rate δ – in order to invest in any asset. The no-arbitrage condition (4) entails that any investment in the safe asset yields (i) a zero (net) profit when there are no transaction costs, and (ii) a negative (net) profit when transactions costs are nonzero. As a consequence, at every date $t \in \{0, 1, 2\}$, agents either invest into shares, or do not invest at all.

3.2. Investment Fund

We label *investment fund* a non-bank financial institution with an open-end structure, that pools liquid resources individually possessed by primary investors, and (re)invests them into a portfolio of profitable, but risky and potentially illiquid assets. Without significant loss of generality, such risky assets can be interpreted as claims on future, stochastic cash flows generated by underlying (real) investment projects. For the sake of simplicity, we assume further that (i) portfolio selection is delegated to a single fund manager, so that decision-making is fully *centralized*, and that (ii) the fund is *not* leveraged. Note that the no-leverage assumption is absolutely realistic, for current financial regulations often forbid the use of leverage to open-ended funds[20]. Moreover, the empirical evidence suggest that closed-end entities, too, make limited use of leverage in practice[21]. Hence, the no-leverage assumption does not significantly impair the possibility to extend of our framework of analysis to closed-ended institutions.

Within our stylized framework, the fund operates over three consecutive periods. In particular, it starts its activity at the initial date $t = 0$, and it is subsequently liquidated – at its market value – at the final date $t = 2$. For the sake of simplicity, we assume that the fund's operations follow a well-specified sequential structure, defined as follows. At the initial date $t = 0$, the fund manager issues a unitary mass of new shares, that are immediately offered to the primary investors of cluster F at the initial price per-share P_0, normalized to

[20]In the US, open- and closed-end vehicles are subject to SEC registration, and are regulated under the Investment Company Act of 1940, the Securities Act of 1933, and the Securities Exchange Act of 1934 – see Giannetti and Kahraman (2017).

[21]Albeit the same financial regulations permit the use of leverage to closed-ended entities – see again Giannetti and Kahraman (2017).

one[22]. Each primary investor is allowed to purchase at most one single share. No additional shares are issued in subsequent dates $t \in \{1, 2\}$. The fund manager invests all financial resources collected via the primary emission into a portfolio of risky assets, that generates a stream of (stochastic) cash flows over an infinite time horizon. The first cash flow v_0 realizes instantaneously in $t = 0$, and every primary investor $i \in [0, 1]$ observes private, noisy information about it. Each primary investor is then called to decide whether to (i) redeem her share at the subsequent (interim) date $t = 1$, or to (ii) wait until the natural maturity of her investments – i.e. until the liquidation of the fund in $t = 2$. In both cases, shares are redeemed at their instantaneous market price, so that (early) redemption in $t = 1$ yields the market price P_1, while liquidation in $t = 2$ yields the market price P_2. No dividends are paid during the entire lifetime of the fund, so that the *net* expected return from purchasing one share in $t = 0$ *coincides* with the future (expected) capital gain, realized either in $t = 1$ or $t = 2$ – at the prevailing, instantaneous market conditions[23]. The fund's regulation imposes a *redemption notice period* to primary investors, whereby requests for the early redemption of shares must be submitted *in advance* – see subsection 3.3 for an extensive discussion of the issue. Formally, every primary investor that opts for (early) redemption in $t = 1$ must notify her decision to the fund manager in $t = 0$. As a consequence, *both* liquidation (market) prices P_1 and P_2 are *unknown* to the primary investor at the moment she chooses her course of action. At the interim date $t = 1$, after all requests for early redemption have been collected, the fund manager offers the corresponding mass of shares in the secondary market, and transactions are subsequently settled at the market-clearing price P_1. At the final date $t = 2$ the fund is finally liquidated, and all the outstanding shares are redeemed at their instantaneous market price P_2.

3.3. Redemption Notice Period

Equity capital providers of open-ended vehicles are subject to several regulatory restrictions that significantly limit their ability to liquidate their investments on short notice[24]. Two common restrictions are *redemption notice periods* and *payout periods*[25]. The former mandate that all investors submit their requests for early redemption of shares to the fund

[22]The normalization $P_0 = 1$ simplifies the analysis and entails no significant loss of generality.

[23]In other words, the intrinsic value of shares stems from the underlying pro-rata claims on the future market value of the fund's portfolio of assets.

[24]In opposition to *closed-end* investment schemes, that allow individual investors to dispose freely of their shares via autonomous trade on secondary markets. European VCTs (Venture Capital Trusts), American BDCs (Business Development Companies), Japanese ITs (Investment Trusts), and Australian LICs (Listed Investment Companies) are all examples of closed-end vehicles.

[25]Hedge funds are allowed to impose even more stringent limitations to their investors' ability to redeem shares.

manager *in advance*, while the latter restrict the redemption of shares to predetermined contractual dates – see e.g. Edwards (1999) and Giannetti and Kahraman (2017). Note that both types of regulatory restrictions involve a delay between the date at which investors decide to redeem their shares, and the date at which such shares are materially liquidated by the fund manager. Within our framework, we focus on *redemption notice periods*, and assume that (i) all shares redeemed before the natural maturity of the fund are sold in the secondary market at the interim date $t = 1$, but (ii) investors must notify their decisions to opt for early redemption at the initial date $t = 0$. The fund manager collects all requests for early redemption at the end of $t = 0$, and subsequently places a unique selling order in the secondary market at the interim date $t = 1$. As discussed in subsection 3.2, a direct consequence of the presence of a redemption notice period is that neither the interim market price P_1, nor the final liquidation price P_2 are known at the moment primary investors decide their courses of action. Every action available involves some risk (here in the form of pure price risk). Note that this is openly consistent with the actual functioning of open-ended investment schemes. Chen et al, (2010), for instance, highlight that the Net Asset Value – henceforth, the fund's *NAV* –x that determines liquidation prices for investors of mutual funds is calculated on the basis of the same-day market prices of the underlying portfolio of securities[26].

3.4. Portfolio Composition and Net Asset Value

Recall from subsection 3.2 that, while the investment fund operates only for three consecutive periods – indexed by $t = 0, 1, 2$ –, the assets in its portfolio generate cash flows over an *infinite* time horizon. In order to avoid any ambiguity in the notation, we index the (discrete) timing of cash flows with $\tau \in \{0, 1, 2, \dots\}$, and indicate with v_τ the instantaneous cash flow at date τ. We assume that the ability of an asset to generate cash flows over time is parameterized by its economic fundamentals, the latter being summarized by a single, unidimensional statistic, and constant over time. Recall further that the financial market is assumed efficient, so that market prices reflect, on average, the *true* economic fundamentals of assets, hence the net present values of the – current and future – cash flows they generate. As a consequence, at any date $t \in \{0, 1, 2\}$ the instantaneous market price P_t of one share of the fund must be proportional, on average, to its (instantaneous) Net Asset Value – henceforth, the fund's *NAV* –, defined as the market value/price of its asset holdings net of all non-equity liabilities. Moreover, since the fund is not leveraged by hypothesis, its NAV coincides with the net present value of the stream of – current and future – cash flows

[26]As reported by the authors, funds' NAVs in the US are determined at 4:00 pm by the NASD (National Association of Security Dealers), and disclosed around 6:00 pm.

generated by its asset holdings. For the sake of simplicity, we abstract completely from endogenous portfolio selection, and assume further that the fund manager is able to commit *ex ante* to a conservative investment strategy, aimed at stabilizing the fund's NAV over time. Consistently, we consider a simplified scenario where the composition of the fund's portfolio of assets is (i) exogenous and (ii) constant over time. Besides being convenient for analytical tractability, such assumptions are consistent with both the empirical and anecdotal evidence. The intrinsic financial fragility of open-ended entities encourages fund managers to adopt prudent trading strategies – see e.g. Cherkes et al. (2010) and Giannetti and Kahraman (2017)[27] –, for a bad current performance hinders their future ability to raise new funds – see e.g. Huang et al. (2007) and Spiegel and Zhang (2013)[28].

Formally, we summarize the underlying economic fundamentals of the fund's portfolio of assets into a summary statistic $\theta \in \mathbb{R}$, constant over time by construction. The true fundamentals θ are unknown by all agents in the economy, that share the following informative common prior

$$\theta \sim \mathcal{N}\left(\bar{\theta}, \sigma_\theta^2\right) \, , \tag{5}$$

whit $\bar{\theta} > 0$ the common prior expectation of fundamental θ, and where the variance σ_θ^2 parameterizes the precision of prior information. Without loss of generality, we can interpret the common prior (5) as a summary of all past information publicly available to economic agents at the beginning of the initial date $t = 0$. We define the generic instantaneous cash flow v_τ as

$$v_\tau = \theta + \eta_\tau, \qquad \forall \, \tau = 0, 1, 2, \dots \tag{6}$$

with θ the fundamental component, and where the infinite sequence of instantaneous shocks η_τ is governed by a stochastic process that satisfies the following conditions

$$\eta_\tau \sim \mathcal{N}\left(0, \sigma_\eta^2\right) \qquad \forall \, \tau = 0, 1, 2, \dots \tag{7a}$$

$$\eta_\tau, \eta_{\tau'} \;\; \text{independent for} \;\; \tau \neq \tau', \, \tau = 0, 1, 2, \dots \tag{7b}$$

[27]Cherkes et al. (2010) shows that open-ended vehicles hold on average more liquid securities, while closed-end counterparties tend to trade in more illiquid assets. Giannetti and Kahraman (2017) document the limited effectiveness of open-ended entities in tackling arbitrage opportunities, in turn providing an empirical verification for the theoretical results of Shleifer and Vishny (1997).

[28]Both papers provide sound empirical evidence that highlights how past performance of open-ended investment funds greatly influences the net volume of new investments. Bad performing vehicles are subject, on average, to high(er) outflows of funds, while vehicles with good track-records do not experience significant funding issues.

Note that condition (7b), in turn, entails that

$$\mathbb{E}\left[\,\eta_\tau\,\eta_{\tau'}\,\right] \,=\, 0 \qquad \forall\,\tau \neq \tau'\;, \tag{8}$$

so that conditions (7a)-(7b) jointly guarantee that all instantaneous shocks are i.i.d., hence the sequence $\{\eta_\tau\}_{\tau=0}^{+\infty}$ is a (Gaussian) independent white noise process. In words, definition (6) states that, at any date $\tau = 0, 1, 2, \ldots$, the instantaneous cash flow v_τ generated by the fund's portfolio is determined in part by the actual economic fundamentals of the underlying assets – fundamental component θ –, and in part by exogenous random disturbances – temporary shocks η_τ. We can interpret the purely temporary random components η_τ as summary statistics for all unmodeled shocks at the macro and/or at the industry level, that induce irregular periodic oscillation around the (unobserved) constant fundamental value θ. We indicate with V_t the *true* fundamental value of (a unitary mass of) the fund's portfolio of assets at date $t \in \{0, 1, 2\}$, defined as the net present value (at date t) of all current and future cash flows v_τ. Formally, we define the (unitary) fundamental value V_t as

$$V_t \,=\, v_t + \sum_{\tau=1}^{+\infty} \Lambda^\tau v_{t+\tau}\;, \tag{9}$$

where v_τ is the instantaneous cash flow at date τ as defined in (6), and with Λ the common discount factor defined in (4). Indicate with $Q_t \in [0, 1]$ the mass of outstanding shares at date t. Since the fund is *not* leveraged, at every date $t \in \{0, 1, 2\}$ its *true* instantaneous NAV is immediately defined as

$$NAV_t \,\equiv\, V_t\big(P_0\,Q_t\big)\;, \tag{10}$$

for every $t \in \{0, 1, 2\}$. In words, at any date t, the fund's NAV coincides with the market value of a unitary mass of its underlying portfolio of asset, multiplied by the instantaneous size $P_0\,Q_t$ of the fund's investment in the portfolio at date t[29]. Since the initial price per-share P_0 is normalized to one, then $P_0\,Q_0 = 1$, and the NAV of the fund at the initial date $t = 0$ is simply $NAV_0 = \theta$. Recall finally that a direct implication of the efficiency of the financial market is that any instantaneous market price per-share P_t must be proportional to the fund's true (unobserved) NAV, divided by the instantaneous mass of outstanding shares

[29] Recall that the initial price P_0 is normalized to one, and that, upon redemption, shares are liquidated at their market price. As a consequence, at any date $t \in \{0, 1, 2\}$ the instantaneous mass Q_t of outstanding shares coincides with the total amount of financial resources still available to the fund at that date. Hence, it parameterizes the instantaneous 'size' of the fund's investment.

Q_t. Since from definition (10) we have that

$$\frac{NAV_t}{Q_t} \equiv V_t \ , \tag{11}$$

then, at any date $\in \{0, 1, 2\}$, the instantaneous market price per-share P_t must be proportional to the unitary fundamental value V_t of the underlying portfolio of assets – i.e. $P_t = P(V_t)$ for every $t \in \{0, 1, 2\}$. As a consequence, the equity nature of primary investors' claims constitutes an implicit insurance against *dilution*, for individual redemptions do not affect the market valuation of outstanding shares directly[30]. Throughout the paper, we indicate with P_1^* and P_2^* the *equilibrium* market prices at the interim and final date, respectively.

3.5. Primary Investors

Primary investors are the atomistic market agents randomly assigned to cluster F. These investors are active in the primary market, i.e. they potentially buy one share of the investment fund at the initial price $P_0 = 1$. The investment is potentially profitable, but intrinsically risky, as its return is unknown *ex ante*, and potentially negative *ex post*. Upon purchasing the share, a primary investor becomes a shareholder of the fund, and she is faced with a binary decision: (i) wait until the contractual maturity of her investment – i.e. until the liquidation of the fund in $t = 2$ –, or (ii) ask for the redemption of the share at the interim date $t = 1$. Formally, we indicate with $a_0^i = 0$ the decision to wait until $t = 2$, and with $a_0^i = 1$ the decision to ask for early redemption in $t = 1$, so that the individual action sets can be written as

$$\mathcal{A}_0^i = \{0, 1\} \ ,$$

for every $i \in [0, 1]$, with a_0^i the generic element. Since investors rely on need to fund their investment decisions through external financing, the debts of all primary investors that choose to wait until $t = 2$ must be rolled over in $t = 1$. The *ex post* sequences of action-contingent wealth endowments $w_t^i(a_0^i)$ of the generic i-th primary investor can be simply defined as:

$$w_t^i \Bigg(\underbrace{a_0^i = 0}_{\text{wait}} \Bigg) = \begin{cases} 0 & \text{in } t = 0 \\ 0 & \text{in } t = 1 \\ P_2^* - (1 + c)(1 + \delta)^2 & \text{in } t = 2 \end{cases} \tag{12}$$

[30] Although they *indirectly* affect instantaneous market prices via their signaling effect on potential buyers – see Section 4.

in case she decides to wait until the final liquidation date $t = 2$, and

$$w_t^i \left(\underbrace{a_0^i = 1}_{\text{redeem}} \right) = \begin{cases} 0 & \text{in } t = 0 \\ P_1^* - (1+c)(1+\delta) & \text{in } t = 1 \\ 0 & \text{in } t = 2 \end{cases} \tag{13}$$

in case she opts for early redemption in $t = 1$. P_1^* and P_2^* the equilibrium market prices at the interim ad final date, respectively. Recall from subsections 3.2 and 3.3 that, due to the presence of a redemption notice period, early redemptions are settled in $t = 1$ but must be notified to the fund manager in $t = 0$, hence neither P_1^* nor P_2^* are known at the moment primary investors make their decisions. As a consequence, both courses of action are intrinsically *risky*. Once notified, individual redemption decisions are *irreversible*. Before choosing optimally their courses of action, primary investors privately learn about the cash flow v_0 generated by the fund's portfolio in $t = 0$. Their private information is summarized by noisy, unbiased signal x_0^i, defined as follows

$$x_0^i = v_0 + \varepsilon_0^i \ , \tag{14}$$

for all $i \in [0, 1]$, with v_0 the instantaneous cash flow realized at date $t = 0$ – see definition (6) –, and where ε_0^i is an idiosyncratic noise term, defined as

$$\varepsilon_0^i \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \ , \tag{15}$$

for all $i \in [0, 1]$. Noise variables ε_0^i are assumed i.i.d. and independent from the sequence of random disturbances η_τ of cash flows – see definitions (7a)-(7b) – and from the fundamental variable θ – see definition (5). In other words, assumption 15 posits that x_0^i represents an unbiased signal on the true value the cash flow in $t = 0$. Hence, x_0^i is indirectly informative about θ. As a consequence, upon observing private information x_0^i primary investors are able to form meaningful and consistent expectations about both future (equilibrium) market prices P_1^* and P_2^*.

3.6. Secondary Market

The secondary market is a (small) niche of the financial market where the shares emitted *ex novo* by the fund in $t = 0$, and subsequently redeemed by primary investors, are offered by the fund manager to secondary investors. Within this market, the aggregate supply $O_1(v_0) \in [0, 1]$ coincides with the total mass of shares redeemed by primary investors in

$t = 0$ – i.e., formally

$$O_1(v_0) \equiv \int_0^1 a_0^i \, \mathrm{d}i \ , \tag{16}$$

with a_0^i the individual selling decision of the i-th primary investor at date $t = 0$. Note that,since the aggregate volume of early redemptions is determined at the initial date $t = 0$, i.e. before a market price P_1 is formed, the aggregate market supply $O_1(v_0)$ defined in (16) is *completely inelastic* to price. Formally, we have that

$$\frac{\partial}{\partial P_1} O_1(v_0) = 0 \tag{17}$$

The mass $s \geq 1$ of atomistic secondary investors populates the demand side of the market. After observing noisy information about the aggregate supply $O_1(v_0)$, each secondary investor can either issue ($a_1^j = 1$) or not ($a_1^j = 0$) a purchasing order for a single share at the posted price P_1. The secondary market is cleared via a Walrasian tâtonnement process, whereby a market auctioneer (i) calls a price P_1, then (ii) elicits purchasing orders from secondary investors *at that price*, and (iii) iterates the procedure until the posted price clears the market. Purchasing orders are not settled until the process determines a market-clearing price P_1^*. Since price discovery arises from a Walrasian auction, secondary investors need not form expectations about the market-clearing price P_1^*: if the posted price does not clear the market, individual purchasing orders are not settled, and can be modified at the subsequent iterations of the process in light of the new posted prices. Due to the continuum-player specification of the market game, for any arbitrary price P_1 – and any realization of the aggregate supply $O_1(v_0)$ – the aggregate demand $B_1(P_1; O_1)$ can be defined as

$$B_1\big(P_1; O_1\big) \equiv \int_0^s a_1^j \, \mathrm{d}j \ , \tag{18}$$

where $O_1(v_0)$ is the aggregate market supply defined in (16). Once determined, the unique market-clearing price P_1^* is observed with no noise by all agents in the economy.

Secondary investors. Secondary investors are the atomistic market agents randomly assigned to cluster S. At the interim date $t = 1$, they participate into the secondary market as potential buyers of the securities sold to liquidate the shares redeemed in $t = 0$ by primary investors. Upon purchasing one share in $t = 1$, a secondary investor is entitled to receive the final liquidation price per share P_2^* at the subsequent (final) date $t = 2$. Since no dividends are paid, the individual gross pay-off *ex post* coincides with the realized capital gain (loss). Within the Walrasian auction mechanism that governs transactions in the secondary market, every j-th secondary investor either issues ($a_1^j = 1$) or not ($a_1^j = 0$) a purchasing order for one single share at the price called by the auctioneer. Transactions are not settled until the market clears. The action set of a generic secondary investor can be defined as

$$\mathcal{A}_1^j = \{0, 1\} \ ,$$

for every $j \in [0, s]$, with a_1^j the generic element. Once the market-clearing price P_1^* is determined, every secondary investor that issued a purchasing order at that price must settle the transaction by borrowing an amount $P_1^* (1 + c)$ of financial resources. The *ex post* sequence of wealth endowments $w_t^j(a_1^j)$ induced by the purchase of one share by the generic j-th secondary investor is

$$w_t^j \Bigg(\underbrace{a_1^j = 1}_{\text{purchase}} \Bigg) = \begin{cases} 0 & \text{in } t = 0 \\ 0 & \text{in } t = 1 \\ P_2^* - P_1^* (1 + c) (1 + \delta) & \text{in } t = 2 \end{cases} \tag{19}$$

with P_1^* the equilibrium clearing price of the secondary market at the interim date $t = 1$, and P_2^* the (equilibrium) liquidation price per-share at the final date $t = 2$. Before the Walrasian auction begins, secondary investors receive private information about the aggregate market supply of shares $O_1 (v_0)$. Building on Dasgupta (2007), we assume that secondary investors' private information is summarized by noisy, unbiased signals z_1^j in the form

$$z_1^j = \Phi^{-1}\big(O_1\big) - \xi_1^j \ , \tag{20}$$

for every $j \in [0, s]$, with $\Phi^{-1} (\cdot)$ the inverse of the Normal Standard CDF[31], and where ξ_1^j is an idiosyncratic noise term defined as follows

$$\xi_1^j \sim \mathcal{N}\left(0, \sigma_\xi^2\right) \ , \tag{21}$$

for every secondary investor $j \in [0, s]$. Recall from definition (16) that the aggregate supply of shares coincides with the volume of early redemptions, and the latter are in turn determined by the rational behavior of privately informed primary investors – whose private information is summarized by private signals x_0^i, defined by expressions (14) and (15). Hence, the market supply $O_1 (v_0)$ is (indirectly) informative about the unknown fundamentals θ of the fund's

[31] Note that the function $\Phi^{-1} (\cdot)$ maps the interval $(0, 1)$ onto the real line, so that $z_1^j \in \mathbb{R}$ even if $O_1 (v_0) \in [0, 1]$. Moreover, the function $\Phi^{-1} (\cdot)$ is strictly increasing in its argument, so that (i) signals z_1^j are strictly increasing in the realized volume of early redemptions $O_1 (v_0)$, and (ii) every realized volume $O_1 (v_0)$ is unambiguously mapped into a single, real value.

portfolio of asset, that determine the final liquidation price P_2^* of shares.

Residual market agents. Residual market agents do *not* directly participate in trade into the secondary market. However, they are allowed to observe the market-clearing price P_1^* that emerged from all transactions settled in the secondary market. Upon observing the market price P_1^*, every residual agent is required, at $t = 2$, to publicly communicate her estimate about the fundamental value of the fund's portfolio of assets. We indicate with $P_2^\ell(P_1^*)$ the estimate communicated in $t = 2$ by the generic ℓ-th residual agent, conditional to the clearing price P_1^*. As discussed in subsection 3.4, at any date $t \in \{0, 1, 2\}$ the instantaneous NAV of the fund coincides with its (instantaneous) fundamental value V_t – see (9) and (10). We can therefore define the generic assessment $P_2^\ell(P_1^*)$ as

$$P_2^\ell(P_1^*) \equiv \mathbb{E}\Big[V_2 \,\Big|\, P_1^* \Big] , \tag{22}$$

for every $\ell \in [0, r]$. In light of our assumption that residual agents do *not* possess private information, disagreement is ruled out *a priori*, and individual assessments must necessarily *coincide*. Formally, it must hold that

$$P_2^\ell(P_1^*) = P_2^{\ell'}(P_1^*) = P_2(P_1^*) , \qquad \forall\, \ell \neq \ell' \tag{23}$$

for every pair of residual agents $\ell, \ell' \in [0, r]$. As such, throughout the paper we focus on the generic ℓ-th residual agent, and consistently omit the individual-specific superscript. In particular, we can define the generic market assessment $P_2(P_1^*)$ at the final date $t = 2$ as

$$P_2(P_1^*) \equiv \left(\frac{1}{1-\Lambda}\right) \mathbb{E}\Big[\theta \,\Big|\, P_1^* \Big] + \sum_{\tau=0}^{+\infty} \Lambda^\tau\, \mathbb{E}\Big[\eta_{2+\tau} \,\Big|\, P_1^* \Big] . \tag{24}$$

Expression (24) clearly highlights that the *current* market-clearing price P_1^* directly affects the *future* liquidation price P_2. Hence, the individual early redemptions notified at the initial date $t = 0$ directly affect both the market-clearing price P_1^* and the final liquidation price P_2.

3.7. Dynamic Structure and Equilibrium

All the elements outlined in the previous subsections can be summarized into a sequential game $\Gamma(\theta)$ with the following dynamic structure:

Initial period $t = 0$ begins

0.a) the fund issues a unitary mass of new shares, and offers one share to each primary investor $i \in [0, 1]$ at the initial price $P_0 = 1$;

0.b) the mass of collected financial resources is invested into a portfolio of risky assets, whose long-term profitability is parameterized by their unobserved economic fundamentals $\theta \in \mathbb{R}$;

0.c) the first instantaneous cash flow v_0 realizes;

0.d) every primary investor privately observes a noisy signal x_0^i, informative about v_0, and decides whether to (i) redeem her share at the (subsequent) interim date $t = 1$ (formally, $a_0^i = 1$), or to (ii) wait until the final liquidation of the fund in $t = 2$ (formally, $a_0^i = 0$);

0.e) period $t = 0$ ends.

Interim period $t = 1$ begins

1.a) the fund manager places a unique selling order in the secondary market for the aggregate mass $O_1(v_0)$ of shares redeemed by primary investors;

1.b) every secondary investor $j \in [0, s]$ learns about the aggregate supply of shares $O_1(v_0)$ via a private, noisy signal z_1^j;

1.c) a market auctioneer calls a price P_1, and every secondary investor decides whether or not ot issue a purchasing order for one share at that price;

1.d) the Walrasian tâtonnement process is iterated until a price P_1^* clears the market;

1.e) period $t = 1$ ends.

Final period $t = 2$ begins

2.a) all residual agents $\ell \in [0, r]$ observe the market-clearing price P_1^*, and publicly announce their (common) estimate P_2^* of the value per-share of the fund's portfolio of assets;

2.b) the fund ceases to operate, and all the outstanding shares are liquidated at their market price P_2^*;

2.c) the game ends.

An equilibrium of the sequential game $\Gamma(\theta)$ is defined as follows.

DEFINITION. An equilibrium of the sequential Bayesian game $\Gamma(\theta)$ consists of: (**I**) a profile $\mathbf{a_0}$ of monotone strategies $a_0^i : \mathbb{R} \longmapsto \{0, 1\}$ for primary investors, with a system of beliefs $\mu_F(\theta|X_0) : \mathbb{R} \longmapsto [0, 1]$; (**II**) a profile $\mathbf{a_1}$ of monotone strategies $a_1^j : \mathbb{R} \longmapsto \{0, 1\}$ for secondary investors, with a system of beliefs $\mu_S(\omega|Z_1) : \mathbb{R} \longmapsto [0, 1]$; (**III**) a price

$P_1^* \in \mathbb{R}_+$ for the secondary market; and (**IV**) a profile $\mathbf{a_2}$ of strategies $a_2^\ell : \mathbb{R} \longmapsto \mathbb{R}$ for residual market agents, with a system of beliefs $\mu_R(\theta | \ln P_1^*) : \mathbb{R} \longmapsto [0,1]$; such that

(i) $a_0^i(x_0^i) \in \arg\max_{a_0^i \in \{0,1\}} a_0^i \, \mathbb{E}[P_1^* \,|\, x_0^i] + (1 - a_0^i)\Big(\Lambda \, \mathbb{E}[P_2^* \,|\, x_0^i]\Big)$, for every primary investor $i \in [0,1]$;

(iii) the secondary market clears at P_1^*, i.e. $B_1\Big(\ln P_1^*; O_1(v_0)\Big) = O_1(v_0)$;

(iv) $a_2^\ell(\ln P_1^*) = \mathbb{E}[V_2 \,|\, \ln P_1^*]$, for every residual market agent $\ell \in [0,r]$, where V_2 is the fundamental value defined in (9);

(v) all conditional posterior beliefs $\mu_F(\theta \,|\, X_0)$, $\mu_S(\theta \,|\, Z_1)$, and $\mu_R(\theta \,|\, \ln P_1^*)$ are pinned down by Bayes rule.

4. Equilibrium Characterization

In this Section we fully characterize the unique monotone equilibrium of the dynamic market game $\Gamma(\theta)$, outlined in Section 3. Given the sequential nature of $\Gamma(\theta)$, we proceed backwards from the last subgame – i.e. from the market game at the final date $t = 2$ where the residual agents of cluster R determine the liquidation price per-share $p_2^*(p_1)$. For the sake of analytical tractability, we specify the market game in terms of *log prices*, and indicate with p_t the natural logarithm of the instantaneous price P_t at the generic date t.

4.1. Final Liquidation Price

As outlined in subsection 3.6, the fund's (log) liquidation price per-share at the final date $t = 2$ is determined by residual market agents. As the secondary market for shares is assumed informationally efficient, p_1 is akin to an informative signal that aggregates all the decentralized information publicly and privately available to individual traders. Within this framework, and in the spirit of Amador and Weill (2010)[32], rational agents can *learn from prices* – see also Angeletos et al. (2006, 2018) and Goldstein et al. (2011, 2013). Note that the aggregation technology that governs the informational content of p_1 – hence the extent of the corresponding rational learning process – is in turn determined by the optimal behavior of traders, hence it is entirely *endogenous*. In order to be able to proceed with the analysis, we need to impose additional structure onto the information aggregation technology that governs the informational content of market prices. We do so by 'guessing' a functional form for the relation that links the observed market-clearing price p_1 to the unobserved economic

[32] It is worth highlighting that the interpretation of (competitive) market prices as informative signals about unobserved underlying economic fundamentals traces back to Hayek (1945).

fundamentals θ, and then proceed by verifying that our guess is indeed correct in equilibrium. In particular, we assume henceforth that:

A.1 The (log) market-clearing price p_1 is a *linear* function of the unobserved cash flow v_0. Hence, it can be expressed as a (linear) function of (i) the fundamental θ and (ii) the instantaneous, non-fundamental disturbance η_0. Formally, we impose that

$$p_1 \;=\; p_1\left(\theta, \eta_0\right) \;=\; \bar{\psi} + \psi_\theta\, \theta + \psi_\eta\, \eta_0 \;, \tag{25}$$

where coefficients $\bar{\psi}, \psi_\theta, \psi_\eta \in \mathbb{R}$ are endogenous and uniquely determined in equilibrium.

Note that, in equilibrium, the aggregation technology (25) is common knowledge. As a consequence, at the final date $t = 2$ all agents in the economy are able to perform model-consistent statistical inference upon observing the market price p_1. In particular, expression (25) can easily be rewritten as

$$\frac{p_1 - \bar{\psi}}{\psi_\theta} \;=\; \theta + \left(\frac{\psi_\eta}{\psi_\theta}\right) \eta_0 \;. \tag{26}$$

Direct inspection of (26) reveals that: (i) the LHS is a linear function of the observed price p_1 with *known* coefficients $\bar{\psi}$ and ψ_θ; (ii) the RHS is a simple sum of the unobserved fundamental θ with a zero-mean noise variable η_0, the latter scaled by a *known* coefficient ψ_η/ψ_θ. For the sake of simplicity, we rewrite the LHS of expression (26) as

$$\frac{p_1 - \bar{\psi}}{\psi_\theta} \;\equiv\; Y_2\left(p_1\right) \;\; , \tag{27}$$

and the scaling coefficient ψ_η/ψ_θ of the instantaneous disturbance η_0 as

$$\frac{\psi_\eta}{\psi_\theta} \;\equiv\; \kappa \;, \tag{28}$$

so that expression (26) can be rewritten as

$$Y_2\left(p_1\right) \;=\; \theta + \kappa\, \eta_0 \;. \tag{29}$$

Expression (29) reveals that the element $Y_2\left(p_1\right)$ constitutes, *de facto*, an unbiased (noisy) signal, informative about the unobserved fundamental θ, whose probability distribution – conditional to θ – is

$$Y_2\left(p_1\right) \,\Big|\, \theta \;\sim\; \mathcal{N}\Big(\, \theta \,,\, \kappa^2 \sigma_\eta^2 \Big) \;. \tag{30}$$

where κ is the endogenous coefficient defined in (28). It is immediate to notice from (30) that the intrinsic precision of the price signal $Y_2(p_1)$ is indeed endogenous, as it directly depends on the the endogenously determined equilibrium coefficients ψ_θ and ψ_η. Given the normality of the common prior (5) and of the conditional distribution of $Y_2(p_1)$ – see expression (30) –, the posterior beliefs for θ of the generic ℓ-th residual agent, conditional to the endogenous signal $Y_2(p_1)$ can be written as

$$\theta \,|\, Y_2(p_1) \;\sim\; \mathcal{N}\Bigg(\alpha\, Y_2(p_1) + (1-\alpha)\,\bar{\theta}\,,\; (1-\alpha)\,\sigma_\theta^2 \Bigg)\,, \tag{31}$$

with $\bar{\theta}$ and σ_θ^2 the common prior expectation and variance of θ, respectively, and where the coefficient $\alpha \in (0,1)$ is defined as follows

$$\alpha \;=\; \frac{\sigma_\theta^2}{\sigma_\theta^2 + \kappa^2 \sigma_\eta^2}\;. \tag{32}$$

Note that p_1 is also indirectly informative about the instantaneous shock η_0. However, considering the assumption that the zero-mean random shocks η_τ to instantaneous cash flows v_τ are i.i.d. and have no persistence, η_0 is not informative about future shocks η_τ – with $\tau \geq 1$. Formally, we have that

$$\mathbb{E}\left[\,\eta_\tau \,\middle|\, p_1\,\right] \;=\; \mathbb{E}\left[\,\eta_\tau\,\right] \;=\; 0\,, \qquad \forall\, \tau \geq 1. \tag{33}$$

Substituting for (33) into definition (24) we are able to characterize the equilibrium final liquidation price $p_2^*(p_1)$, expressed as a function of the clearing price p_1 of the secondary market at the interim date $t = 1$. We summarize the result into the following Lemma.

Lemma 4.1. *The equilibrium liquidation price per-share $p_2^*(p_1)$ at the final date $t = 2$, expressed as a function of the market-clearing price p_1 of the secondary market, is*

$$p_2^*(p_1) \;=\; \left(\frac{1}{1-\Lambda}\right)\left[\alpha\left(\frac{p_1 - \bar{\psi}}{\psi_\theta}\right) + (1-\alpha)\,\bar{\theta}\right] \tag{34}$$

with $\bar{\theta}$ the common prior expectation of the unobserved fundamental θ, $\Lambda \in (0,1)$ the (common) discount factor defined in (4), $\alpha \in (0,1)$ the expectation coefficient defined in (32), and where $\bar{\psi}$ and ψ_θ are the endogenous coefficients of the market-clearing price p_1.

A direct implication of Lemma 4.1 is that a *temporary* disturbance η_0 to the initial cash flow v_0 still affects the (equilibrium) market price per-share p_2^* at the final date $t = 2$ – i.e. two periods after its occurrence. Interestingly enough, the result holds true notwithstanding the informational efficiency of the financial market. Even more surprising is the fact that the result holds true *because of* the informational efficiency of the financial market. What generates such an unduly persistence of the pecuniary effects of a purely temporary, non-fundamental shock is indeed the endogenous learning process that pins down the optimal decision-making of market agents. This holds because η_0 directly affects the early redemption of shares by primary investors via private information x_0^i, hence the aggregate supply of shares $O_1(v_0)$ in the secondary market. Notice further that also secondary investors' decisions are affected by the temporary shock: they extract information about the unknown fundamental θ from the (private) noisy observation of the aggregate market supply $O_1(v_0)$. Finally, due to the informational efficiency of the financial market, all private information possessed by secondary investors is aggregated by the clearing price p_1 and publicly disclosed to residual agents, that (efficiently) use it to determine the final price $p_2^*(p_1)$. In other words, the same informational efficiency that allows for the transmission of valuable information via market prices (potentially) acts as a propagation mechanism for non-fundamental shocks that affect the transmitted information *ab origine*. It is worth noting that our result is fully consistent with both the theoretical prescriptions of the 'limited arbitrage theory' of Shelifer and Vishny (1997), and the related empirical literature. The latter, in particular, provides robust evidence that the investment flows directed to open-ended vehicles are heavily influenced by the *past* performance of the latter – on the issue, see e.g. Huang et al. (2007) and Spiegel and Zhang (2013).

4.2. Clearing Price of the Secondary Market

Having characterized the unique final equilibrium price per-share $p_2^*(p_1)$, we are now able to determine the (unique) equilibrium price p_1^* that clears the secondary market for shares. Similarly to subsection 4.1, we adopt a 'guess and solve' approach, and begin our analysis by stating a sensible working hypothesis about the equilibrium behavior of primary investors.

A.2 All primary investors $i \in [0,1]$ adopt a *monotone* equilibrium strategy, whereby a request for early redemption is issued to the fund manager at the initial date $t = 0$ if (and only if) private information x_0^i is sufficiently bad. Formally, we assume that, in equilibrium, every primary investor behaves according to the following strategy in cutoff form

$$a_0^i = 1 \quad \Longleftrightarrow \quad x_0^i \leq \hat{x} \ , \tag{35}$$

for every $i \in [0,1]$, where $\hat{x} \in \mathbb{R}$ is an *arbitrary* cutoff value common to all primary investors and *common knowledge* in equilibrium.

The monotone strategy (35) essentially prescribes that a primary investor opts for the early redemption of her share if (and only if) she is sufficiently pessimistic about the quality of the fund's portfolio of assets. Indeed, when a primary investor estimates that the fundamentals of the fund are bad, she expects that *both* the equilibrium (interim) market-clearing price p_1^* and the final liquidation price p_2^* will be low. But a low clearing price p_1^* is in turn equivalent to the public disclosure of bad news to the market, due to the informational efficiency of the latter. As a consequence, a low clearing price p_1^* observed at the interim date $t = 1$ is subjectively expected to result into an even lower liquidation price p_2^* at the final date $t = 2$. A direct implication of such an amplification mechanism is that it is rational for a pessimistic primary investor to sell her share as soon as possible, hence the only optimal course of action is to ask for early redemption.

4.2.1. Learning From the Market Supply

If all primary investors indeed comply with the prescriptions of the monotone strategy described by expression (35), the aggregate mass $O_1(v_0; \hat{x})$ of shares redeemed in $t = 1$ can be defined as

$$O_1(v_0; \hat{x}) \;=\; \Pr\left(x_0^i \leq \hat{x} \,\middle|\, v_0 \right) \,, \tag{36}$$

due to the continuum-player nature of the game [33], with v_0 the cash flow realized at the initial date $t = 0$. Given expression (36) and the formal definition (6) of cash flow v_0, we can formally define the aggregate market supply $O_1(v_0; \hat{x})$ as

$$O_1(v_0; \hat{x}) \;=\; \Phi\left(\frac{\hat{x} - \theta - \eta_0}{\sigma_\varepsilon} \right) \,, \tag{37}$$

where σ_ε is the standard deviation common to instantaneous shocks ε_0^i. Considering that secondary investors are privately informed about the market supply $O_1(v_0; \hat{x})$ via the observation of noisy signals z_0^j we can substitute (37) into definition (20), and with some trivial algebraic manipulations, we obtain

$$\hat{x} - \sigma_\varepsilon \, z_1^j \;=\; \theta + \eta_0 + \sigma_\varepsilon \, \xi_1^j \;. \tag{38}$$

The LHS of expression (38) is a linear function of the observed signal z_1^j, with *known* coefficients[34] $\hat{x}$ and σ_ε. Since all future cash flows v_1, v_2, ... are *not* affected by the instantaneous

[33] For the sake of formal rigor, the equivalence highlighted by expression (36) holds *almost surely*. In words, since the financial system under analysis is populated by infinitely many atomistic agents, the actual mass of primary investors with a private signal $x_0^i \leq \hat{x}$ coincides (almost surely) with the true probability that the event $x_0^i \leq \hat{x}$ occurs, conditional to the realized cash flow v_0.

[34] Recall indeed that the unique cutoff value $\hat{x}$ is common knowledge in equilibrium.

shock η_0, the latter essentially constitutes *aggregate noise*. Hence, the RHS of expression (38) is a simple sum of the unobserved fundamental θ with a non-fundamental, zero-mean shock $(\eta_0 + \sigma_\varepsilon\, \xi_1^j)$, where (i) η_0 is an aggregate noise component, common to all secondary investors, and (ii) ξ_1^j is an *idiosyncratic noise* component, specific to the j-th (secondary) investor. Notice further that the idiosyncratic noise component ξ_1^j is scaled by the (known) standard deviation σ_ε of primary investors' private information. Given the formal definition (37) of the aggregate market supply $O_1\,(v_0; \hat{x})$, and the technology (20) that governs signals z_i^j, the interpretation is straightforward. As the precision of primary investors' private information x_0^i increases, the informational content of their (aggregate) actions about the unobserved cash flow v_0 increases, too. In other words, the magnitude of the signaling effect embedded into primary investors' selling decisions increases in the precision of the private information that pins down such decisions. For the sake of simplicity, we rewrite the LHS of expression (38) as

$$\hat{x} - \sigma_\varepsilon\, z_1^j \;=\; Z_1^j\,(\hat{x}) \;\; . \tag{39}$$

It is immediate to check from the expression (38), that the element $Z_1^j\,(\hat{x})$ is an unbiased, noisy (private) signal, informative about the unobserved fundamental θ, with conditional distribution

$$Z_1^j\,(\hat{x})\;|\,\theta \;\sim\; \mathcal{N}\Big(\,\theta\,,\,\sigma_\eta^2 + \sigma_\varepsilon^2\sigma_\xi^2\,\Big)\;\; , \tag{40}$$

for every $j \in [0, s]$ and every $\hat{x} \in \mathbb{R}$, where σ_ξ^2 is the variance common to all idiosyncratic noise variables ξ_1^j – see definition (21) –, and with

$$Cov\,\Big(\,Z_1^j\,(\hat{x})\,, Z_1^{j'}\,(\hat{x})\;\big|\,\theta\,\Big) \;=\; \sigma_\eta^2\;\; , \tag{41}$$

for every $j \neq j'$, with $j, j' \in [0, s]$. Note from (41) that the presence of an aggregate random disturbance η_0 induces (conditional) correlation among the endogenous private signals $Z_1^j\,(\hat{x})$ of secondary investors [35].. It is finally worth highlighting that definition (39) entails that the endogenous signals Z_1^j strictly increase in the cutoff value $\hat{x}$ of primary investors' monotone strategies. Formally, we have that

$$\frac{\partial}{\partial\,\hat{x}}\,Z_1^j\,(\hat{x}) \;=\; 1 \;>\; 0\;\; . \tag{42}$$

The interpretation is again intuitive. When the equilibrium cutoff value $\hat{x}$ is very low, only primary investors that received (very) bad news about v_0 opt for early redemption. As a consequence, for every *ex post* volume of early redemptions $O_1\,(v_0; \hat{x})$, a low(er) cutoff $\hat{x}$

[35] However, similarly to primary investors' private information x_0^i, the endogenous signals Z_1^j are *uncorrelated* conditioning on the (unobserved) cash flow v_0.

signals, on average, a low(er) quality θ of the underlying assets. An increase in the cutoff value $\hat{x}$ prescribes that, in equilibrium, primary investors with better signals opt for early redemption, too. As a consequence, the quality θ of the underlying assets is higher, on average, for every *ex post* volume of early redemptions $O_1(v_0; \hat{x})$. A direct implication of such mechanics is that, upon observing – with some noise – the *same* aggregate mass of early redemptions $O_1(v_0; \hat{x})$, secondary investors are (on average) more optimistic about the unobserved quality θ of the assets in portfolio, as the cutoff value $\hat{x}$ of primary investors increases.

4.2.2. Optimal Purchasing Rule

Since shares by assumption do *not* pay any dividend at any date, at the interim date $t = 1$ a secondary investor is willing to borrow money for the purchase one share if she expects a net capital gain at the final date $t = 2$. In other words, secondary investors need to form expectations about the future liquidation price $p_2^*(p_1)$ in order to decide rationally whether or not to issue a purchasing order at the posted price p_1. Moreover, secondary investors know that the future liquidation price $p_2^*(p_1)$ will be proportional to the unobserved fundamentals θ of the fund's portfolio of assets, in reason of the efficiency of the financial market. As a consequence, at the interim date $t = 1$ their subjective expectations $p_2^*(p_1)$ are pinned down by their subjective (conditional) expectations about θ. Recall from expression (40) that all private signals $Z_1^j(\hat{x})$ of secondary investors are (conditionally) normally distributed, hence the posterior beliefs of a generic secondary investor can be defined as

$$\theta \,|\, Z_1^j(\hat{x}) \;\sim\; \mathcal{N}\left(\beta\, Z_1^j(\hat{x}) + (1-\beta)\,\bar{\theta}\,,\; (1-\beta)\,\sigma_\theta^2 \right) \,, \tag{43}$$

and

$$\eta_0 \,|\, Z_1^j(\hat{x}) \;\sim\; \mathcal{N}\left(\gamma \left(Z_1^j(\hat{x}) - \bar{\theta} \right) \,,\; \gamma \left(\sigma_\theta^2 + \sigma_\varepsilon^2 \sigma_\xi^2 \right) \right) \,, \tag{44}$$

with

$$Cov\left(\theta, \eta_0 \,\middle|\, Z_1^j(\hat{x}) \right) \;=\; -\beta\, \sigma_\eta^2 \;=\; -\gamma\, \sigma_\theta^2 \,, \tag{45}$$

for every $i \in [0, s]$ – where the coefficients β and γ are defined as

$$\beta \;=\; \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2} \,, \tag{46}$$

and

$$\gamma \;=\; \frac{\sigma_\eta^2}{\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2} \,. \tag{47}$$

Recall further from subsection 3.6 that the secondary market clears via a Walrasian tâtonnement process, whereby a market auctioneer calls a price p_1 and elicits purchasing orders at that price until the price-contingent demand $B_1(p_1; O_1)$ equates the price-inelastic supply $O_1(v_0; \hat{x})$. As a consequence, secondary investors need not form expectations about the (log) market-clearing price p_1^*. Finally, recall from expression (1) that, at any date $t \in \{0, 1, 2\}$, the instantaneous utility of a market agent coincides with her (instantaneous) financial wealth, and that the no-arbitrage condition (4) entails that $\Lambda(1+\delta) = 1$. It is therefore immediate to check from (19) that, for any posted (log) price p_1, it is rational for a secondary investor to issue a purchasing order if – and only if

$$\Lambda\,\mathbb{E}\left[\exp\{p_2^*(p_1)\}\,\middle|\,Z_1^j(\hat{x})\right] \;\geq\; (1+c)\Big(\exp\{p_1\}\Big)\,, \tag{48}$$

with $p_2^*(p_1)$ the equilibrium liquidation price defined in Lemma 4.1. Moreover, from expression (25) we know that $p_2^*(p_1)$ is (conditionally) normally distributed, hence the conditional expectation at the LHS of expression (48) can be written as

$$\mathbb{E}\left[\exp\{p_2^*(p_1)\}\,\middle|\,Z_1^j(\hat{x})\right] \;=\; \exp\left\{\mathbb{E}\left[p_2^*(p_1)\,\middle|Z_1^j(\hat{x})\right] + \frac{1}{2}\,Var\left(p_2^*(p_1)\,\middle|Z_1^j(\hat{x})\right)\right\}\,. \tag{49}$$

Finally, note from the definition (29) of the endogenous price signal $Y_2(p_1)$ that the conditional expectation of the equilibrium final liquidation price $p_2^*(p_1)$ defined in Lemma 4.1 can be written as

$$\mathbb{E}\left[p_2^*(p_1)\,\middle|\,Z_1^j(\hat{x})\right] \;=\; \left(\frac{\alpha}{1-\Lambda}\right)\left(\mathbb{E}\left[\,\theta\,\middle|\,Z_1^j(\hat{x})\right] + \kappa\,\mathbb{E}\left[\,\eta_0\,\middle|\,Z_1^j(\hat{x})\right]\right) + \left(\frac{1-\alpha}{1-\Lambda}\right)\bar{\theta}\,, \tag{50}$$

with $\bar{\theta}$ the common prior expectation of the unknown fundamental θ, and where κ is the scaling coefficient of the price signal defined in (28) and α is the expectation coefficient defined in (32). Moreover, the conditional variance $\varSigma^*$ of secondary investors' estimates in $t = 1$ of the equilibrium final liquidation price $p_2^*(p_1)$ is defined as

$$\begin{aligned}\varSigma^* \;\equiv\; Var\Big(p_2^*(p_1)\,\Big|\,Z_1^j(\hat{x})\Big) \;=\; \left(\frac{\alpha}{1-\Lambda}\right)^2 \Bigg(& Var\left(\,\theta\,\middle|\,Z_1^j(\hat{x})\right) \;+ \\ &+ \kappa^2\,Var\left(\,\eta_0\,\middle|\,Z_1^j(\hat{x})\right) \;-\; 2\,\kappa\,Cov\left(\,\theta, \eta_0\,\middle|\,Z_1^j(\hat{x})\right)\Bigg)\,.\end{aligned} \tag{51}$$

A proper substitution of the conditional posterior beliefs (43)-(47) in expressions (50) and (51) unambiguously determines the conditional expectation and variance of $p_2^*(p_1)$ of the generic j-th secondary investor at the interim date $t = 1$. Substituting further into the rational purchasing rule (48) we obtain that, for any arbitrary price p_1 posted by the Walrasian auctioneer, it is optimal for a secondary investor to place a purchasing order if – and only if

$$p_1 \;\leq\; \hat{p}_1\left(Z_1^j\left(\hat{x}\right)\right) \;, \tag{52}$$

where the critical value $\hat{p}_1\left(Z_1^j\left(\hat{x}\right)\right)$ is

$$\hat{p}_1\left(Z_1^j\left(\hat{x}\right)\right) \;=\; \bar{\Pi}^* + \Pi_Z^*\, Z_1^j\left(\hat{x}\right) \;, \tag{53}$$

with $\bar{\Pi}^*$ and Π_Z^* the two endogenous coefficients defined as

$$\bar{\Pi}^* \;=\; \left(\frac{\sigma_\theta^2\left(\sigma_\eta^2\left(1-\kappa\right)+\sigma_\varepsilon^2\sigma_\xi^2\right)+\kappa^2\sigma_\eta^2\left(\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)}{\left(1-\Lambda\right)\left(\sigma_\theta^2+\kappa^2\sigma_\eta^2\right)\left(\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)}\right)\bar{\theta}+\frac{1}{2}\,\Sigma-\ln\left(\frac{1+c}{\Lambda}\right) \tag{54}$$

and

$$\Pi_Z^* \;=\; \frac{\sigma_\theta^2\left(\sigma_\theta^2+\kappa\,\sigma_\eta^2\right)}{\left(1-\Lambda\right)\left(\sigma_\theta^2+\kappa^2\sigma_\eta^2\right)\left(\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)} \;, \tag{55}$$

respectively, and where Σ is the conditional variance of secondary investors' estimates at the interim date $t = 1$ about the equilibrium liquidation price $p_2^*(p_1)$ at the final date $t = 2$. The latter is defined as follows

$$\Sigma \;=\; \left(\frac{\sigma_\theta^2}{\left(1-\Lambda\right)\left(\sigma_\theta^2+\kappa^2\sigma_\eta^2\right)}\right)^2\left(\frac{\sigma_\theta^2\left(\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)+\kappa^2\sigma_\eta^2\left(\sigma_\theta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)-2\,\kappa\,\sigma_\theta^2\sigma_\eta^2}{\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2}\right) \;. \tag{56}$$

In words, the optimal purchasing rule (52) states that it is rational for a secondary investor to place a purchasing order for one single share of the fund if (and only if) the posted price p_1 is sufficiently low – where the endogenous cutoff value for purchase $\hat{p}_1$ directly depends on the investor's private information $Z_1^j\left(\hat{x}\right)$. We summarize the result into the following Lemma.

Lemma 4.2. *In equilibrium, all secondary investors $j \in [0, s]$ place their purchasing orders according to the following monotone strategy*

$$a_1^j \;=\; 1 \qquad \Longleftrightarrow \qquad p_1 \;\leq \hat{p}_1\left(Z_1^j\left(\hat{x}\right)\right) \;,$$

for any arbitrary monotone strategy $\hat{x} \in \mathbb{R}$ adopted by primary investors, where the critical value $\hat{p}_1\left(Z_1^j\left(\hat{x}\right)\right)$ is defined by expression (53).

Lemma 4.2 essentially states that a secondary investor is willing to place a purchasing order if (and only if) the information she extracts from the observed market supply suggests that the future (log) liquidation price $p_2^*(p_1)$ will be high enough to warrant a total (future) expenditure of $p_1(1+c)(1+\delta)$. Note from expression (53) that, before we fully characterize the equilibrium clearing price p_1^*, it is impossible to determine whether the equilibrium cutoff price for purchase increases or decreases in private information $Z_1^j(\hat{x})$. Indeed, differentiating $\hat{p}_1\left(Z_1^j(\hat{x})\right)$ with respect to $Z_1^j(\hat{x})$ we obtain

$$\frac{\partial}{\partial\, Z_1^j(\hat{x})}\,\hat{p}_1\left(Z_1^j(\hat{x})\right) \;=\; \Pi_Z^* \;, \tag{57}$$

for every $j \in [0, s]$, where the sign of Π_Z^* depends, in turn, on the sign of the endogenous coefficient κ defined in (28), unambiguously determined in equilibrium but whose sign is *ex ante* unknown[36]. It is however sensible to *guess* that the cutoff price $\hat{p}_1\left(Z_1^j(\hat{x})\right)$ is *strictly increasing* in private information. The interpretation is straightforward. Since the endogenous signals $Z_1^j(\hat{x})$ strictly decrease in the aggregate volume $O_1(v_0;\hat{x})$ of early redemptions, to every secondary investor an high(er) observed signal is akin to good news. As a consequence, as $Z_1^j(\hat{x})$ increases investors' subjective estimates of the unobserved fundamental θ are revised upwards, and so it must be, *ceteris paribus*, their willingness to pay for the purchase of one share. In Section 5 we formally prove that both our guess and the proposed economic interpretation are indeed correct.

4.2.3. Market Clearing

Transactions on the secondary market are not settled until the posted price clears the market. A direct inspection of the formal definition (53) of the (subjectively) optimal cutoff value for purchase $\hat{p}\left(Z_1^j(\hat{x})\right)$ immediately reveals that the equilibrium purchasing rule of Lemma 4.2 can be rewritten as

$$a_1^j \;=\; 1 \qquad \Longleftrightarrow \qquad Z_1^j(\hat{x}) \;\geq\; \hat{Z}_1(p_1) \;\;, \tag{58}$$

[36]From the foraml definition (55) of Π_Z^* it is indeed immediate to check that

$$\kappa \;<\; -\frac{\sigma_\theta^2}{\sigma_\eta^2} \;<\; 0 \quad\Longrightarrow\quad \Pi_Z^* \;<\; 0 \;,$$

where κ is the endogenous coefficient defined in (28).

where the price-contingent cutoff value $\hat{Z}_1(p_1)$ is defined as

$$\hat{Z}_1(p_1) = \frac{p_1 - \bar{\Pi}^*}{\Pi_Z^*} , \tag{59}$$

and with $\bar{\Pi}^*$ and Π_Z^* the endogenous coefficients defined in (54)-(56). In reason of the continuum-player specification of the game under analysis, at any posted price p_1 the total mass of purchasing orders coincides (almost surely) with the probability of the event $Z_1^j(\hat{x}) \geq \hat{Z}_1(p_1)$ – conditional to the unobserved cash flow v_0. As for to the cutoff price $\hat{p}_1\left(Z_1^j(\hat{x})\right)$, we are not yet able to determine whether the critical value $\hat{Z}_1(p_1)$ for purchase increases or decreases in the posted price p_1, as it is impossible to determine *ex ante* the sign of the endogenous price coefficient Π_Z^*. Recall from definition (37) that the cutoff value $\hat{x}$ that governs the equilibrium behavior of primary investors determines the aggregate supply of shares $O_1(v_0; \hat{x})$ in the secondary market. Hence, all the non-pecuniary (informational) effects exerted by $O_1(v_0; \hat{x})$ on the aggregate demand via the private signals $Z_1^j(\hat{x})$ observed by secondary investors are fully summarized by the cutoff value $\hat{x}$. The aggregate demand can therefore be written as $B_1(p_1; O_1) = B_1(p_1; \hat{x})$, and formally defined as follows

$$B_1(p_1; \hat{x}) = s\left[\Phi\left(\frac{\theta + \eta_0 - \hat{Z}_1(p_1)}{\sigma_\varepsilon \sigma_\xi}\right)\right] , \tag{60}$$

with $s \geq 1$ the total mass of secondary investors, and where $\hat{Z}_1(p_1)$ is the optimal cutoff defined in (59). Note that, similarly to Hellwig et al. (2006) and Goldstein et al. (2013), both aggregate market demand $B_1(p_1; \hat{x})$ and the aggregate market supply $O_1(v_0; \hat{x})$ take the form of cumulative distribution functions[37]. Using expression (60) for the aggregate demand, and expression (37) for the aggregate supply, the market-clearing condition for the secondary market can be written as

$$\underbrace{\Phi\left(\frac{\hat{x} - \theta - \eta_0}{\sigma_\varepsilon}\right)}_{O_1(v_0)} = \underbrace{s\left[\Phi\left(\frac{\theta + \eta_0 - \hat{Z}_1(p_1)}{\sigma_\varepsilon \sigma_\xi}\right)\right]}_{B_1(p_1; \hat{x})} . \tag{61}$$

For the sake of analytical tractability, we assume henceforth that $s = 1$, i.e. we focus on

[37] In Goldstein et al. (2013) the aggregate market demand takes the form of a normal standard CDF as a consequence of the optimal behavior of market agents. However, building on Hellwig et al. (2006) and Dasgupta (2007), the authors introduce the *ad hoc* assumption that the aggregate market supply has the same functional form for the sake of analytical tractability. On the contrary, in our model the functional form $\Phi(\cdot)$ is *fully micro-founded* for both the market demand and the market supply.

the specific scenario where the total mass of potential buyers of shares coincides with the total mass of (potential) sellers. It is immediate to notice from definition (60) that, for any posted price p_1, the mass of purchasing orders strictly increases in s, hence the latter can easily be interpreted as (a proxy for) market-wide liquidity. An extensive discussion of the issue is deferred to Section 5. Solving for p_1 the market-clearing condition (61) we obtain that the unique equilibrium price $p_1^* (\theta, \eta_0; \hat{x})$ that clears the secondary market for shares is

$$p_1^* (\theta, \eta_0; \hat{x}) \;=\; \Pi_Z^* \, (1 + \sigma_\xi) \, (\theta + \eta_0) - \sigma_\xi \, \Pi_Z^* \, \hat{x} + \bar{\Pi}^* \ , \tag{62}$$

with Π_Z^* and $\bar{\Pi}^*$ the endogenous coefficients defined in (54)-(56), and where $\hat{x}$ is the cutoff value of primary investors' monotone strategies. By matching the coefficients of definition (62) with those of the arbitrary guess (25) it is immediate to notice that

$$\psi_\theta^* \;=\; \psi_\eta^* \;=\; \Pi_Z^* \, (1 + \sigma_\xi) \ \ , \tag{63}$$

and

$$\bar{\psi}^* \;=\; \bar{\Pi}^* - \sigma_\xi \, \Pi_Z^* \hat{x} \ . \tag{64}$$

Recall that $v_0 = \theta + \eta_0$, hence the equilibrium market-clearing price can be conveniently expressed as a function of the cash flow realized at the initial date $t = 0$. Moreover, as a consequence of (63), we have that

$$\kappa \;=\; \frac{\psi_\eta}{\psi_\theta} \;=\; 1 \ . \tag{65}$$

Once the equilibrium value of the endogenous coefficient $\kappa = 1$ is determined, so are all the endogenous coefficients. As a consequence, for any arbitrary cutoff value $\hat{x}$ adopted by primary investors for their monotone strategies at $t = 0$, the unique equilibrium price $p_1^* (v_0; \hat{x})$ that clears the secondary market is unambiguously determined. We summarize the result into the following Lemma.

Lemma 4.3. *For every $s \geq 1$ and every $\hat{x} \in \mathbb{R}$, there exists a unique equilibrium price $p_1^* (v_0; \hat{x})$ that clears the secondary market for shares at the interim date $t = 1$. When $s = 1$, the equilibrium market-clearing price is*

$$p_1^* (v_0; \hat{x}) \;=\; \left(\frac{1}{1 - \Lambda} \right) \Big(\, \bar{\Psi}^* \, (\hat{x}) + \Psi_v^* \, v_0 \Big) \ , \tag{66}$$

with $\Lambda = 1 \, (1 + \delta)$ the common discount factor, and where the equilibrium coefficients $\bar{\Psi}^* \, (\hat{x})$ and Ψ_v^* are defined as

$$\Psi_v^* \; = \; \frac{\sigma_\theta^2 \left(1 + \sigma_\xi\right)}{\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2} \tag{67}$$

and

$$\bar{\Psi}^* \left(\hat{x}\right) \; = \; \left(\frac{1}{\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2}\right) \left(\left(\sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2\right) \bar{\theta} - \left(\sigma_\xi \, \sigma_\theta^2\right) \hat{x}\right) + \frac{1}{2} \, \Sigma^* - \ln \left(\frac{1+c}{\Lambda}\right) \; , \tag{68}$$

with $c > 0$ the per-unit transaction cost, and where the element Σ^* is defined as

$$\Sigma^* \; = \; \frac{\left(1 - \Lambda\right) \sigma_\theta^4 \sigma_\varepsilon^2 \sigma_\xi^2}{\left(\sigma_\theta^2 + \sigma_\eta^2\right) \left(\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2 \sigma_\xi^2\right)} \; . \tag{69}$$

Lemma 4.3 highlights two interesting results. First, the equilibrium market-clearing price $p_1^* \left(v_0; \hat{x}\right)$ is strictly increasing in the underlying economic fundamental θ, for the endogenous coefficient Ψ_v^* is strictly positive for every arbitrary strategy $\hat{x}$ adopted by primary investors and every calibration of the exogenous parameters. Hence, the true economic fundamentals of assets indeed parameterize their instantaneous market value. Second, the equilibrium price $p_1^* \left(v_0; \hat{x}\right)$ constitutes *de facto* an *unbiased*, noisy signal, informative about the unknown asset fundamentals θ. Hence, the secondary market for shares is indeed efficient, from both an allocative and an informational perspective. Note however that, similarly to the final liquidation price $p_2^* \left(p_1\right)$ characterized in Lemma 4.1, the market-clearing price $p_1^* \left(v_0; \hat{x}\right)$ is directly affected by the purely temporary innovation η_0 that only affects the first cash flow v_0. The result seems to confirm the hypothesis whereby endogenous learning process induced by the signaling effect of market price(s) contributes significantly to the inter-temporal propagation of temporary, non-fundamental shocks.

Having fully characterized the (interim) equilibrium price $p_1^* \left(v_0\right)$ of the secondary market for shares, we are now able to determine the (unique) equilibrium price per-share p_2^* at the final date $t = 2$. Recall indeed that Lemma 4.1 characterizes p_2^* as a function of the (generic) clearing price p_1. We summarize the result into the following Lemma.

Lemma 4.4. *When $s = 1$, the equilibrium liquidation price per-share $p_2^* \left(v_0\right)$ at the final date $t = 2$ is defined as*

$$p_2^* \left(v_0\right) \; \equiv \; p_2^* \left(p_1^*\right) \; = \; \left(\frac{1}{1 - \Lambda}\right) \left[\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\eta^2}\right) v_0 + \left(\frac{\sigma_\eta^2}{\sigma_\theta^2 + \sigma_\eta^2}\right) \bar{\theta}\right] \; , \tag{70}$$

with $\Lambda \; = \; 1/\left(1 + \delta\right)$ the common discount factor and $\bar{\theta}$ the common prior expectation of *fundamental θ, and where v_0 is the instantaneous cash flow realized at the initial date $t = 0$.*

Note finally from expression (70) that, differently from the equilibrium market-clearing price p_1^*, the (equilibrium) final price p_2^* does *not* depend on the strategy $\hat{x}$ adopted by primary investors to redeem shares in $t = 0$.

4.3. Early Redemptions by Primary Investors

At the end of the initial date $t = 0$, every primary investor $i \in [0, 1]$ is called to decide whether to redeem her share (i.e. $a_0^i = 1$) or to wait until the liquidation of the investment fund at the final date $t = 2$ (i.e. $a_0^i = 0$). Investors that choose to wait until $t = 2$ are liquidated at the instantaneous market price $P_2^*(v_0)$ prevailing at the final date $t = 2$. Investors that opt for early redemption are liquidated at the market price $P_1^*(v_0; \hat{x})$ prevailing at the interim date $t = 1$. Before making their decision, all primary investors are privately informed about the first realized cash flow v_0 via the noisy signals x_0^i defined in (14)-(15). Based on their private information, primary investors will choose the course of action that maximizes their expected utilities. As such, we can argue that early redemption in $t = 0$ is subjectively optimal if – and only if

$$\mathbb{E}\Big[\, P_1^*(v_0; \hat{x}) \;\Big|\; x_0^i \Big] \;\geq\; \Lambda\, \mathbb{E}\Big[\, P_2^*(v_0) \;\Big|\; x_0^i \Big] \,. \tag{71}$$

Upon observing her private information x_0^i, the generic i-th primary investor holds the following set of conditional posterior beliefs

$$v_0 \,|\, x_0^i \;\sim\; \mathcal{N}\Big(\, \varphi\, x_0^i + (1 - \varphi)\, \bar{\theta} \,,\; \varphi\, \sigma_\varepsilon^2 \Big) \,, \tag{72}$$

for every $i \in [0, 1]$, with v_0 the cash flow realized at date $t = 0$, and where the expectation coefficient φ is defined as

$$\varphi \;=\; \frac{\sigma_\theta^2 + \sigma_\eta^2}{\sigma_\theta^2 + \sigma_\eta^2 + \sigma_\varepsilon^2} \,. \tag{73}$$

Via the initial guess (**A.2**) we assume that, in equilibrium, all primary investors choose whether or not to redeem their shares in $t = 0$ according to the monotone strategy (35). The latter prescribes to opt for early redemption when $x_0^i \leq \hat{x}$, i.e. when private information suggests that the quality θ of the fund's asset portfolio might be poor. Considering that any monotone strategy with arbitrary switching at $\hat{x}$ has to be compatible with the optimality condition (71), it must hold that

$$x_0^i \le \hat{x} \quad \Longrightarrow \quad \mathbb{E}\Big[\, P_1^*\left(v_0;\hat{x}\right) \,\big|\, x_0^i \,\Big] \,\ge\, \Lambda\, \mathbb{E}\Big[\, P_2^*\left(v_0\right) \,\big|\, x_0^i \,\Big] \,. \tag{74}$$

It is immediate to check from Lemma (4.3) and (4.4) that primary investors' conditional expectations for both the equilibrium market-clearing price $p_1^*\left(v_0;\hat{x}\right)$ and the final (equilibrium) liquidation price $p_2^*\left(v_0\right)$ are (i) linear and (ii) strictly increasing in private information x_0^i, for any arbitrary cutoff value $\hat{x}$. In particular, we have that

$$\frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_1^*\left(v_0;\hat{x}\right) \,\big|\, x_0^i \,\Big] \;=\; \left(\frac{1}{1-\Lambda}\right) \left(\frac{\sigma_\theta^2\left(\sigma_\theta^2+\sigma_\eta^2\right)\left(1+\sigma_\xi\right)}{\left(\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\right)\left(\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2\sigma_\xi^2\right)}\right) \tag{75}$$

and

$$\frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_2^*\left(v_0\right) \,\big|\, x_0^i \,\Big] \;=\; \left(\frac{1}{1-\Lambda}\right) \left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\eta^2+\sigma_\varepsilon^2}\right) \,. \tag{76}$$

When the two slopes are *not* identical, there exists a unique value $\hat{x}$ where the two conditional expectations intersect. Hence, a *necessary* condition for both the existence and the uniqueness of an equilibrium monotone strategy of the type outlined in (35) is

$$\frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_2^*\left(v_0\right) \,\big|\, x_0^i \,\Big] \;\neq\; \frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_1^*\left(v_0;\hat{x}\right) \,\big|\, x_0^i \,\Big] \,. \tag{77}$$

Moreover, for such unique monotone strategy to be consistent with the optimal behavior prescribed by expression (74), the linear conditional expectations for the final equilibrium price $p_2^*\left(v_0\right)$ must be *steeper* in x_0^i than the conditional expectation for the interim market-clearing price $p_1^*\left(v_0;\hat{x}\right)$ for all primary investors. Otherwise, the (unique) monotone strategy with switching at $\hat{x}$ would prescribe that (i) all investors with private information x_0^i such that $\Lambda\mathbb{E}\left[p_2^*\left(v_0\right)|x_0^i\right] > \mathbb{E}\left[p_1^*\left(v_0;\hat{x}\right)|x_0^i\right]$ opt for the early redemption of their shares in $t=0$, while (ii) all investors with private information x_0^i such that $\mathbb{E}\left[p_1^*\left(v_0;\hat{x}\right)|x_0^i\right] > \Lambda\mathbb{E}\left[p_2^*\left(v_0\right)|x_0^i\right]$ choose to wait until the final liquidation of the fund in $t=2$. In both cases, the behavior prescribed by the monotone strategy $\hat{x}$ would be openly suboptimal, hence incompatible with equilibrium. To ensure that an equilibrium monotone strategy $\hat{x}$ indeed exists, we therefore impose the following *consistency condition*

$$\frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_2^*\left(v_0\right) \,\big|\, x_0^i \,\Big] \;>\; \frac{\partial}{\partial\, x_0^i}\, \mathbb{E}\Big[\, p_1^*\left(v_0;\hat{x}\right) \,\big|\, x_0^i \,\Big] \,. \tag{78}$$

Note that if the above consistency condition (78) holds, then the necessary condition for existence (77) is satisfied *a fortiori*. Hence, condition (78) is *sufficient* for the existence of a unique monotone strategy for primary investors in equilibrium. Substituting for the slopes (75) and (76) into expression (78) we can (re)state the consistency condition as follows.

CONSISTENCY CONDITION.

The exogenous precision $1/\sigma_\varepsilon^2$ of primary investors' private information x_0^i is such that

$$\frac{1}{\sigma_\varepsilon^2} < \frac{\sigma_\xi}{\sigma_\theta^2 + \sigma_\eta^2} \ , \tag{79}$$

for every $i \in [0,1]$, with $\left(\sigma_\theta^2 + \sigma_\eta^2\right)^{-1}$ the precision of the common prior for v_0, and where σ_ξ parameterizes the precision of secondary investors' private information.

The consistency condition essentially imposes that the precision of the information privately observed by primary investors be not too precise with respect to precision of public information available *ex ante* to all other market agents[38]. In other words, we are imposing that the informational advantage of primary investors over the other atomistic market agents be somehow limited. Note that the constraint imposed by condition (79) onto primary investors' private information is proportional to (a function of) the precision $1/\sigma_\xi^2$ of the private information available to secondary investors. Indeed, as the latter increases, the constraint imposed by the consistency condition becomes more stringent.

When the consistency condition (79) holds, it is immediate to check that, in order to be compatible with equilibrium, the unique monotone strategy $\hat{x}$ that governs primary investors' redemptions at date $t = 0$ must satisfy the following (indifference) condition

$$\mathbb{E}\Big[p_1^*(v_0;\hat{x}) \ \big| \ \hat{x}_0^i = \hat{x}\Big] \ = \ \mathbb{E}\Big[p_2^*(v_0) \ \big| \ x_0^i = \hat{x}\Big] \ , \tag{80}$$

otherwise $\hat{x}$ would prescribe a suboptimal behavior to *some* primary investors. For expressions (66) and (68) of Lemma 4.3 we know that all conditional expectation(s) of $p_1^*(v_0;\hat{x})$ are linear and strictly decreasing in in the cutoff value $\hat{x}$[39]. As a consequence, there exists a unique value $\hat{x} = x_0^*$ that satisfies (80). Solving for the optimal cutoff x_0^* the indifference condition (80) we are finally able to characterize the unique equilibrium monotone strategy of primary investors. We summarize the result into the following Lemma.

[38]Indeed, since the temporary shocks η_0 to the instantaneous cash flows v_0 are i.i.d. and independent of θ, the common prior (unconditional) variance of v_0 is $Var(v_0) = Var(\theta) + Var(\eta_0) = \sigma_\theta^2 + \sigma_\eta^2$.

[39]Graphically, we have that any increase in $\hat{x}$ amounts to a parallel shift downwards of the *entire* expectation mapping $\mathbb{E}\left[p_1^*(v_0;\hat{x})\,|x_0^i\right]$

Lemma 4.5. *Let $s = 1$. Then, in equilibrium, all primary investors $i \in [0,1]$ choose whether or not to redeem their shares in $t = 0$ according to the following monotone strategy*

$$a_0^i = 1 \qquad \Longleftrightarrow \qquad x_0^i \leq x_0^* \ ,$$

with the critical value x_0^* defined as follows

$$x_0^* = \left(\frac{2\,\sigma_\xi}{1 + \sigma_\xi} \right) \bar{\theta} + \chi^* \ , \tag{81}$$

where χ^* is an endogenous coefficient defined as follows

$$\chi^* = \left(\frac{1}{\sigma_\theta^2 \sigma_\xi \left(1 + \sigma_\xi\right)\left(1 - \varphi\right)} \right) \left[\ln\left(\frac{1+c}{\Lambda^\Lambda} \right) + \frac{\sigma_\varepsilon^2\, \varphi}{2\left(1 - \Lambda\right)} \left(\alpha^2 - \left(\Phi_v^*\right)^2 \right) - \beta \left(\frac{\sigma_\theta^2 \sigma_\varepsilon^2 \sigma_\xi^2}{2\left(1 - \Lambda\right)} \right) \right] \ , \tag{82}$$

and with $\bar{\theta}$ the common prior expectation of the fundamental θ.

4.4. Participation Constraint and Equilibrium

In order to complete the characterization of the (unique) monotone equilibrium of the sequential game under analysis, we need to determine under which condition(s) the initial purchase of shares is indeed optimal for primary investors. In case the purchase does *not* occur, the investor experiences a zero total utility, for she has no exogenous financial endowment and investing in the safe asset is always suboptimal due to the no-arbitrage condition (4). In case the purchase indeed occurs, the investor's total utility depends on her future course of action. In case of early redemption she receives $P_1^*\left(v_0\right) - \left(1 + c\right)$ at the interim date $t = 1$. In case she wait until the fund's liquidation, her net utility gain at the final date $t = 2$ is $P_2^*\left(v_0\right)$. Then, the participation constraint at the initial date $t = 0$ can be written as

$$\min \left\{ \Lambda\, \mathbb{E}\Big[P_1^*\left(v_0\right) \Big] \, , \, \Lambda^2\, \mathbb{E}\Big[P_2^*\left(v_0\right) \Big] \right\} \geq 1 + c \ , \tag{83}$$

with $\Lambda = 1/\left(1 + \delta\right)$ the common discount factor, and where $c > 0$ is the sunk (per-unit) transaction cost. Recall from subsection 3.4 that the only information available to primary investors about the fundamental θ, at the moment they participate into the fund's primary emission, is the common prior (5). Absent any private information, their subjective *ex ante* expectations about θ, hence about the future (equilibrium) market prices, must necessarily agree. As a consequence, all primary investors $i \in [0,1]$ have the same participation constraint at $t = 0$. Solving for the condition $\Lambda\, \mathbb{E}\left[P_1^*\left(v_0\right)\right] \geq 1 + c$ we obtain that every primary investor that expects to redeem early her share is willing to purchase at the beginning of date $t = 0$ if – and only if

$$\bar{\theta} \geq \bar{\theta}^* (P_1^*) = (1 - \Lambda) \ln \left(\frac{1 + c}{\Lambda} \right) - \frac{\sigma_\theta^4}{2 (1 - \Lambda) \left(\sigma_\theta^2 + \sigma_\eta^2 \right)} . \tag{84}$$

By the same reasoning, every primary investor that expects to wait until the final liquidation of the fund is willing to purchase at the beginning of date $t = 0$ is – and only if

$$\bar{\theta} \geq \bar{\theta}^* (P_2^*) \tag{85}$$

The participation constraint of primary investors at $t = 0$ can therefore be written as

$$\bar{\theta} \geq \min \left\{ \bar{\theta}^* (P_1^*) , \bar{\theta}^* (P_2^*) \right\} \tag{86}$$

with $\bar{\theta}$ the common prior expectation of the unknown fundamental θ, and where $\bar{\theta}^* (P_1^*)$ and $\bar{\theta}^* (P_2^*)$ are the critical values defined by expressions (84) and (85), respectively. We can now fully characterize the unique monotone equilibrium of our market game. We summarize the result into the following Proposition.

Proposition 1. *Let $s = 1$ and the participation constraint (86) hold. Then, the sequential game $\Gamma(\theta)$ has a unique monotone Bayes-Nash equilibrium $\left\langle x_0^*, \hat{p}_1 (Z_1) , P_1^*, P_2^*, \mu_F (\theta \,|\, X_0) , \mu_S (\theta \,|\, Z_1) , \mu_R (\theta \,|\, \ln P_1^*) \right\rangle$, such that:*

- *(i) at the initial date $t = 0$, all primary investors $i \in [0, 1]$ redeem their shares according to the monotone strategy x_0^* defined in Lemma 4.5;*
- *(ii) at the interim date $t = 1$, all secondary investors $j \in [0, 1]$ place their purchasing orders for shares according to the monotone strategy $\hat{p}_1 \left(Z_1^j (x_0^*) \right)$ defined in Lemma 4.2;*
- *(iii) a unique price per-share $p_1^* (v_0)$, defined as in Lemma 4.3, clears the secondary market;*
- *(iv) at the final date $t = 2$, all residual market agents $\ell \in [0, r]$ announce a final liquidation price per-share $p_2^* (v_0)$, defined as in Lemma 4.4;*
- *(v) all conditional beliefs $\mu_F (\theta \,|\, X_0)$, $\mu_F (\theta \,|\, Z_1)$, and $\mu_F (\theta \,|\, \ln P_1^*)$ are derived via Bayes rule.*

5. Preliminary Results

Building on the equilibrium characterization outlined and extensively discussed in Section 4, in this Section we present and comment three preliminary results. First, in the spirit of Brunnermeier and Pedersen (2009), we characterize the total liquidity of a financial security as the result of the interaction between two separate – but intertwined – components: an (i) asset-specific liquidity component, pinned down by the (unobserved) economic fundamentals of the security; and a (ii) market-wide liquidity component, entirely determined by the prevailing (exogenous) market conditions. We show that, at any date, the instantaneous market price of the security is unambiguously determined by these two liquidity components. Second, we highlight that asset-specific (il)liquidity might induce a significant mispricing of financial assets, even when the market is informationally efficient. Interestingly, asset mispricing might be *amplified* by the informational efficiency of the market, for the noisy information indirectly conveyed by the instantaneous market prices tends to propagate in time the pecuniary effects of purely temporary, non-fundamental shocks. Third, we characterize and discuss the implicit, price-mediated feedback effect whereby asset-specific (il)liquidity directly affects the rational decision-making of atomistic market agents. In particular, we determine under which conditions *strategic complementarity* among primary investors' early redemptions arises *endogenously* from asset-specific (il)liquidity. When this is the case, the latter is, at the same time, a primary source and a direct consequence of the fund's financial fragility.

5.1. Asset-Specific Liquidity and Endogenous Fire-Sale Discounts

Recall that, within our framework, the aggregate behavior of primary investors at the initial date $t = 0$ affects – directly and indirectly – both the market-clearing price $p_1^* (v_0)$ at the interim date $t = 1$ *and* the liquidation price $p_2^* (v_0)$ at the final date $t = 2$. In particular, any increase in the total volume of early redemptions triggers two major effects. First, it induces a corresponding decrease in the interim clearing price of the secondary market via a standard *law-of-demand* argument[40]. Indeed, by construction primary investors' early redemptions in $t = 0$ determine the aggregate supply of shares at the subsequent date $t = 1$. Second, larger volumes of redeemed shares signal to the potential buyers in the secondary market that the unobserved fundamentals θ of the fund's portfolio of assets might be bad. Moreover, the adverse pecuniary effects brought about large(r) initial volumes of early redemptions propagate to the final liquidation price $p_2^* (v_0)$ via the informational content of the interim market-clearing price $p_1^* (v_0)$. We refer to the former externality as the *quantity effect* – of

[40] A symmetric argument applies to a decrease in the aggregate supply.

the total mass of early redemptions –, and to the latter as the *learning effect.*

5.1.1. *Quantity Effect*

In order to study both effects formally, it is convenient to express the equilibrium market-clearing price p_1^* as an explicit function $p_1^*(O_1; s)$ of (i) the aggregate supply of shares $O_1(v_0)$, and of (ii) the total mass $s \geq 1$ of potential buyers. From the market-clearing condition (61), we can write

$$p_1^*(O_1; s) \;=\; \bar{\Pi}^* + \Pi_Z^*(\theta + \eta_0) - \sigma_\varepsilon \sigma_\xi \, \Pi_Z^* \left[\Phi^{-1}\left(\frac{1}{s} \, O_1(v_0) \right) \right] , \tag{87}$$

with $\bar{\Pi}^*$ and Π_Z^* the equilibrium coefficients of secondary investors' optimal purchasing rule (53) – see definitions (54) and (55), respectively –, and where $\Phi^{-1}(\cdot)$ is the inverse of the Normal Standard CDF. Note that for every $O_1(v_0) \in (0,1)$ we have that $(1/s)\, O_1(v_0) \in (0, 1/s) \subseteq (0,1)$ for every $s \geq 1$, hence the equilibrium price $p_1(O_1; s)$ is finite and well-defined. Differentiating expression (87) with respect to the *ex post* aggregate market supply $O_1(v_0)$ we obtain

$$\frac{\partial}{\partial O_1} \, p_1^*(O_1; s) \;=\; -\frac{1}{s} \left(\frac{\sigma_\varepsilon \sigma_\xi \, \Pi_Z^*}{\phi\left(\Phi^{-1}\left(\frac{1}{s} \, O_1(v_0) \right) \right)} \right) \;<\; 0 \;, \tag{88}$$

where $\phi(\cdot)$ indicates the Normal Standard PDF. Expression (88) immediately reveals that the equilibrium market-clearing price p_1^* is *strictly decreasing* in the aggregate supply of shares $O_1(v_0)$. We summarize the result into the following Corollary.

Corollary 1. *The equilibrium market-clearing price $p_1^*(v_0)$ is strictly decreasing in the aggregate market supply of shares $O_1(v_0)$.*

The derivative (88) captures the quantity effect discussed above: it parameterizes the instantaneous decrease in the market-clearing price $p_1^*(v_0)$ triggered by a net increase in the aggregate supply of shares, brought about by a corresponding increase in the volume of early redemptions by primary investors.

5.1.2. *Learning Effect*

In order to determine to what extent the equilibrium clearing price p_1^* is affected by the learning effect of primary investors' early redemptions, we consider a counter-factual scenario where all market agents do *not* learn from prices. When this is the case, the informational content of *any* interim market-clearing price is null. Indicating with ${}_{NL}p_t$ the instantaneous market price at date $t \in \{0, 1, 2\}$ under the no-learning assumption, and via the general definition (24) of the final liquidation price p_2, we have that

$$ {}_{NL}p_2^* \equiv \mathbb{E}\Big[V_2 \,\Big|\, {}_{NL}p_1^* \Big] = \mathbb{E}\big[V_2 \big] = \left(\frac{1}{1-\Lambda} \right) \bar{\theta} \ , \tag{89} $$

with $\bar{\theta}$ the common prior expectation of the unknown economic fundamental θ, and where V_2 is the net present value of all future cash flows generated by the fund's portfolio of assets, evaluated at date $t = 2$ – see definition (9). In words: the final liquidation price under the no-learning assumption is equal to the *prior* (common) expectation about the net present value V_2. Since the common prior expectation $\bar{\theta}$ is common knowledge, then all secondary investor $j \in [0, s]$ share the *common certainty* that ${}_{NL}p_2^*$ is equal to (89), whatever their private information $Z_1^j(\hat{x})$. As a consequence, their decisions at $t = 1$ are *not* affected by their private information, and the only rational(izable) purchasing rule is

$$ {}_{NL}a_1^j = 1 \quad \Longleftrightarrow \quad {}_{NL}p_2^* \geq {}_{NL}p_1 + \ln\left(\frac{1+c}{\Lambda} \right) \ . \tag{90} $$

for every $j \in [0, s]$. Note that if all secondary investors adopt the purchasing rule (90) then the informational content of ${}_{NL}p_1$ is indeed null, hence the final price ${}_{NL}p_2^*$ is indeed defined by (89). A direct consequence of (90) is that, absent any learning from the market, the aggregate demand $B_1({}_{NL}p_1; \hat{x})$ is *infinitely elastic* to the (log) market price ${}_{NL}p_1$, so that it must hold that

$$ {}_{NL}p_1^* = \left(\frac{1}{1-\Lambda} \right) \bar{\theta} - \ln\left(\frac{1+c}{\Lambda} \right) \ , \tag{91} $$

for every aggregate supply $O_1(v_0) \in (0, 1)$. In words, when agents do not learn from the market, *any* volume of redeemed shares is sold in the secondary market at the inelastic price ${}_{NL}p_1^*$ defined by (91). When, instead, agents learn from the market, then private information directly affects the individual purchasing decisions of all secondary investors at $t = 1$, and the market-clearing price p_1 acts as an (efficient) information aggregator. When this is the case, the elasticity of the equilibrium market-clearing price $p_1^*(O_1; s)$ to the aggregate supply of shares $O_1(v_0)$ is defined by expression (88). The counter-factual price differential

$$ LD^*(O_1; s) = p_1^*(O_1; s) -_{NL} p_1^* \tag{92} $$

is therefore an implicit measure of the learning effect of primary investors' early redemptions, for it is the price component that directly accrues to the information that agents 'learn' endogenously from the market. Recall that, due to the informational efficiency of the financial market, residual agents $\ell \in [0, r]$ extract information about θ form the equilibrium market-clearing price $p_1^*(v_0)$, and set accordingly a final liquidation price $p_2^*(p_1^*)$. Hence, the learning effect affects the final price $p_2^*(v_0)$, too, via the interim price $p_1^*(v_0)$. In particular, from Lemma 4.1 we have that

$$\frac{\partial}{\partial O_1} p_2^*\big(p_1^*(O_1; s)\big) = \left(\frac{\alpha}{\Psi_v^*(1-\Lambda)}\right) \frac{\partial}{\partial O_1} p_1^*(O_1; s) < 0 , \tag{93}$$

with α the expectation coefficient (32) and Ψ_v^* the equilibrium price coefficient (67) of Lemma 4.3, and where the derivative $\partial p_1^* / \partial O_1$ is defined by expression (88). As the volume of early redemptions increases, the final liquidation price decreases, due to the (adverse) signaling effect exerted by the market-clearing price $p_1^*(v_0)$. Note however that $p_1^*(v_0)$ explicitly internalizes the informational spillover it induces onto the (equilibrium) final liquidation price $p_2^*(v_0)$, hence the indirect learning effect on $p_2^*(v_0)$ defined by expression (93) is implicitly embedded into the price differential $LD^*(O_1; s)$ defined by (92).

5.2. Market-Wide Liquidity

Recall that, within our analytical framework, $s \geq 1$ indicates the (exogenous) total mass of secondary investors. Since the latter constitute the entire demand side of the secondary market, the statistic s also parameterizes the aggregate mass of potential buyers of shares at the interim date $t = 1$. It is immediate to check from the market-clearing condition (61) that any increase in the the total volume of redeemed shares $O_1(v_0)$ is more easily accommodated by the secondary market as the mass s of potential buyers increases, too. As a consequence, any increase in s results, *ceteris paribus*, into a *higher* market-clearing price. We formally derive the result via the definition (87) of the equilibrium clearing price $p_1^*(O_1; s)$ – see subsection 5.1. Differentiating $p_1^*(O_1; s)$ with respect to s we obtain that

$$\frac{\partial}{\partial s} p_1^*(O_1; s) = \left(\frac{O_1(v_0)}{s^2}\right) \left(\frac{\sigma_\varepsilon \sigma_\xi \, \Pi_Z^*}{\phi\Big(\Phi^{-1}\big(\frac{1}{s} O_1(v_0)\big)\Big)}\right) > 0 , \tag{94}$$

with Π_Z^* the endogenous coefficient defined in (55), and where $\phi(\cdot)$ indicates the Normal Standard PDF. The derivative in (94) is always positive: $\Pi_Z^* > 0$ and $O1(v_0)$ are nonnegative by construction, and $\phi(\cdot) \geq 0$ by definition. Therefore, it seems quite sensible to interpret parameter s as a proxy for *market-wide liquidity*, comparable in spirit to the notions of liquidity adopted by Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009) – where liquidity endogenously depends on the "cash in the market" –, and by Matta and

Perotti (2017) – that define liquidity risk as:

> *. . . "a form of short term, nonfundamental price risk due to the temporary scarcity of cash or arbitrage capital in the market."*

In our paper, market-wide liquidity coexists with asset-specific liquidity. Indeed, the overall liquidity of the assets traded in the secondary market (shares emitted by the fund) is determined in part by non-fundamental, market-wide factors – summarized by the exogenous parameter s –, and in part by intrinsic, fundamental characteristics of the traded assets – parameterized by the environmental variable θ.

5.3. Asset Liquidity and (Endogenous) Strategic Complementarity

In Section 4 we characterize the unique equilibrium of our market game starting from the working hypothesis that primary investors adopt monotone equilibrium strategies of the type defined in (35). For any monotone strategy with arbitrary switching at $\hat{x}$ to be compatible with equilibrium, it must hold that

$$x_0^i \leq \tilde{x} \quad \Longrightarrow \quad EU_0^i\left(a_0^i\!=\!1\right) \geq EU_0^i\left(a_0^i\!=\!0\right) \ ,$$

for every primary investor $i \in [0,1]$. Otherwise the strategy (35) would prescribe a suboptimal equilibrium behavior for *some* primary investors. Lemma 4.5 highlights that, within our analytical framework, there exists a unique critical value x_0^* that satisfies the above optimality condition. Recall from subsection 3.6 that the total volume of early redemptions by primary investors at date $t = 0$ determines the aggregate supply of shares in the secondary market at the subsequent date $t = 1$, hence the instantaneous market price $p_1^*\left(v_0\right)$. Due to the informational efficiency of the financial market, the latter in turn affects – both directly and indirectly – the final liquidation price per-share $p_2^*\left(v_0\right)$. The utility of a primary investor is therefore always affected *ex post* by the aggregate behavior of her fellows, whatever her course of action in equilibrium. As consequence, both waiting and early redemption are intrinsically risky actions, from both a fundamental *and* a strategic perspective. Note that no restrictions are imposed *a priori* onto the nature of such strategic interdependence, for the latter is implicitly determined by the feedback mechanism that links the evolution of prices across periods to the instantaneous aggregate behavior of the atomistic market agents. In this subsection we study under which conditions – if any – early-redemption decisions by primary investors are *strategic complements*, in the sense of Bulow et al. (1985).

Since every primary investor faces a binary choice problem, her incentive to switch either course of action to the other available option is pinned down by the (expected) welfare gain brought about by her switching. Indicate with ΔU_0^i the total *ex post* welfare gain[41], evaluated at the initial date $t = 0$, experienced by the i-th primary investor upon choosing to redeem her share instead of waiting until the liquidation of the fund in $t = 2$. From a direct inspection of expressions (12)-(13) it is immediate to check that the net welfare gain ΔU_0^i can be formally defined as

$$\Delta U_0^i \,=\, P_1^*\left(v_0\right) - \Lambda\, P_2^*\left(v_0\right) \;, \tag{95}$$

with $\Lambda = 1/\left(1+\delta\right)$ the common discount factor. Early redemptions at date $t = 0$ are strategic complements if the individual incentive to redeem increases in the aggregate mass of primary investors that opt for th same course of action. Therefore, we can check for the presence of strategic complementarity by studying under which conditions – if any – the *ex post* net welfare gain ΔU_0^i (strictly) increases in the aggregate volume $O_1\left(v_0\right)$ of early redemptions. We know that the equilibrium market-clearing price at the interim date $t = 1$ strictly decreases in the aggregate supply of shares $O_1\left(v_0\right)$. Hence, we can directly link the net welfare gains of *individual* primary investors to their *aggregate* behavior by express ΔU_0^i as a function of the sole equilibrium price $p_1^*\left(v_0\right)$. To do so, we appeal to the definition (34) of Lemma 4.1, that characterizes the final equilibrium price p_2^* as a function of the (equilibrium) interim market price p_1^*. Using the definition (27) of the implicit price signal $Y_2\left(p_1^*\right)$ observed by residual market agents in $t = 2$, we can easily rewrite expression (95) as $\Delta U_0^i = \Delta U_0^i\left(p_1^*\right)$, i.e. formally

$$\Delta U_0^i\left(p_1^*\right) \,=\, \exp\Big\{\, p_1^*\left(v_0\right) \Big\} - \Lambda \exp\left\{ \left(\frac{1}{1-\Lambda}\right) \left[\left(\frac{\alpha}{\Psi_v^*}\right) p_1^*\left(v_0\right) - \left(\frac{\alpha}{\Psi_v^*}\right) \bar{\Psi}^* + \left(1-\alpha\right) \bar{\theta} \right] \right\} \;, \tag{96}$$

with α the expectation coefficient defined in (32), and where Ψ_v^* and $\bar{\Psi}^*$ are the equilibrium coefficients of $p_1^*\left(v_0\right)$. Differentiating expression (96) with respect to the aggregate volume of early redemptions $O_1\left(v_0\right)$ we obtain the following result.

Proposition 2. *Primary investors' early redemptions in $t = 0$ are strategic complements if – and only if*

$$v_0 \,<\, v_0^* \;, \tag{97}$$

where the critical value v_0^* is defined as follows

$$v_0^* \,=\, \left(\frac{1}{\left(1-\Lambda\right)\Psi_v^* - \alpha}\right) \left[\alpha\, \bar{\theta} - \left(1-\Lambda\right) \left(\bar{\Psi}^* - \ln\left(\frac{\alpha}{\left(1-\Lambda\right)\Psi_v^*}\right)\right)\right] \;, \tag{98}$$

[41] That is, the sum of all the (extra) instantaneous utilities over the (primary) investors' entire, three-period lifetime.

with α the expectation coefficient defined in (32), and where Ψ_v^* and $\bar{\Psi}^*$ are the equilibrium coefficients of the unique price $p_1^*(v_0)$ that clears the secondary market. If $v_0 > v_0^*$, then early redemptions are strategic substitutes.

Proposition 2 states that strategic complementarity arises *endogenously* among primary investors' decisions to redeem early their shares when the unobserved cash flow at $t = 0$ is sufficiently bad. On the contrary, substitutability arises when the initial cash flow is large enough. This result allows to endogenously characterize the nature of the strategic interaction among players. While a full understanding of the rationale of this results is still lacking, intuition can help in providing a first insight. When the initial cash flow is "strong", primary investors will receive on average a good signal. As such, the mass of primary investors that will decide to run will be relatively small. This implies that a marginal increase in the number of early withdrawals will have a bigger (depressive) impact on P_1 than on P_2: investors will have a greater incentive to wait until the natural liquidation of the fund in $t = 2$. On the contrary, when the first cash flow is sufficiently low, investors will receive on average a bad signal. For this reason, a large mass of them will decide to opt for an early withdrawal, leading to a depressed P_1. Nevertheless, a further increase in the number of withdrawal at $t = 1$, will lead to a greater contraction (via the signalling effect on secondary investors) on P_2 than on P_1. This interpretation opens to some considerations on the nature of financial fragility. Indeed, on average, when the fundamentals are bad (and so the first cash flow is low), financial fragility will prevail, considering that strategic choices are complementary. On the contrary, when the fundamental is strong, the financial intermediary is not exposed to runs, since investors' strategies are substitutes. On this consideration, a remark is due: this dynamic holds on average. Nevertheless, since the cut-off is defined in terms of the initial cash flow, and not of the fundamental, we might observe inefficient runs. Indeed, there might be some cases where, despite a solid fundamental, the initial cash flow is impacted by a negative temporary shock. In this case, since the initial cash flow drives all the signals in the game, we will observe a fire sales spiral which drives the intermediary into insolvency, despite the solid fundamentals. In this sense, it would be interesting to investigate the scope for policy initiatives, e.g. in terms of information disclosure or targeted monetary support, to limit such inefficient runs.

6. Conclusion

The analysis carried out in the paper provides insights into the nature of financial fragility for non-bank financial intermediaries. In particular, we show how temporary liquidity needs, possibility induced by temporary negative shocks, can lead to an asset sale spiral, which erodes the values of the portfolio of assets of the fund. This, in turn, might lead to the materialization of insolvency risk for the investment fund. As such, our analysis draws a link between (il)liquidity and insolvency risk. Building on informational efficient financial markets, we are able to endogenously determine the nature of the strategic interaction among investors. This brings some insights into the nature of financial fragility for (non-bank) financial intermediaries. In particular, temporary, non fundamental shocks, can lead to inefficient runs and thus lead an otherwise viable intermediary into insolvency.

In terms of future research agenda, we believe that, building on this observation, the current analysis lays the ground for future policy considerations on how to avoid (or limit) such inefficient runs. Furthermore, the analysis can benefit from a simplification of the characterization of investors and of funding markets (and of the payoffs structures).

References

Allen, F., Carletti, E., Goldstein, I., and Leonello, A. (2018), "Government Guarantees and Financial Stability". *Journal of Economic Theory*, 177, pp. 518-557.

Amador, M., and Weill, P. O. (2010), "Learning from Prices: Public Communication and Welfare". *Journal of political Economy*, 118(5), pp. 866-907.

Angeletos, G. M., and Werning, I. (2006), "Crises and Prices: Information Aggregation, Multiplicity and Volatility". *American Economic Review*, 96, pp. 1720-1736.

Angeletos, G. M., Lorenzoni, G., and Pavan, A. (2018), "Wall Street and the Silicon Valley: A Delicate Interaction". Unpublished.

Broecker, T. (1990), "Credit-Worthiness Tests and Interbank Competition". *Econometrica*, 58(2), pp. 429-452.

Brunnermeier, M., and Petersen, L. (2009), "Market Liquidity and Funding Liquidity". *Review of Financial Studies*, 22, pp. 2201-2238.

Brunnermeier, M. (2009), "Deciphering the Liquidity and Credit Crunch 2007–2008". *Journal of Economic Perspectives*, 23(1), pp. 77-100.

Bryant, J. (1980), "A Model of Reserves, Bank Runs, and Deposit Insurance". *Journal of Banking and Finance*, 4, pp. 335-344.

Carlsson, H. and van Damme, E. (1993), "Global Games and Equilibrium Selection". *Econometrica*, 61(5), pp. 989-1018.

Chen, Q., Goldstein, I., and Jiang, W. (2010), "Payoff Complementarities and Financial Fragility: Evidence from Mutual Funds Outflows". *Journal of Financial Economics*, 97, pp. 239-262.

Cherkes, M., Sagi, J., and Stanton, R. (2008), "A Liquidity-Based Theory of Closed-End Funds". *Review of Financial Studies*, 22, pp. 257-297.

Coval, J., and Stafford, E. (2007), "Asset Fire Sales (and Purchases) in Equity Markets". *Journal of Financial Economics*, 86, pp. 479-512.

Dasgupta, A. (2007), "Coordination and Delay in Global Games". *Journal of Economic Theory*, 134, pp. 195-225.

Demirgüç-Kunt, A., Kane, E., and Laeven, L. (2015), "Deposit Insurance Around the World: A Comprehensive Analysis and Database". *Journal of Financial Stability*, 20, pp. 155-183.

Diamond, D. W., and Dybvig, P. H. (1983), "Bank Runs, Deposit Insurance and Liquidity". *Journal of Political Economy*, 91(3), pp. 401-419.

Diamond, D. W., and Rajan, R. G. (1987), "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking". *Journal of Political Economy*, 109(2), pp. 287-327.

Dimson, E., and Kozerski, C. M. (2002), "Closed-End Funds: A Survey". *Financial Markets, Institutions & Instruments*, 8, pp. 1-41.

Edwards, R. E. (1999), "Hedge Funds and the Collapse of Long-Term Capital Management". *Journal of Economic Perspectives*, 13(2), pp. 189-210.

Eisenbach, T., Keister, T., McAndrews, J., and Yorulmazer, T. (2014), "Stability of Funding Models: An Analytical Framework". *Federal Reserve Bank of New York Economic Policy Review*, 20, pp. 29-47.

Fisher, I. (1933), "The Debt-Deflation Theory of Great Depressions". *Econometrica*, 1(4), pp. 337-357.

Flannery, M. J. (1996), "Financial Crises, Payment System Problems, and Discount Window Lending". *Journal of Money, Credit and Banking*, 28(4), pp. 804-824.

Giannetti, M., and Kahraman, B. (2017), "Open-End Organizational Structures and Limits to Arbitrage". *Review of Financial Studies*, 31(2), pp. 773-810.

Goldstein, I., Ozdenoren, E., and Yuan, K. (2011), "Learning and Complementarities in Speculative Attacks". *Review of Economic Studies*, 78, pp. 263-292.

Goldstein, I., Ozdenoren, E., and Yuan, K. (2013), "Trading Frenzies and Their Impact on Real Investment". *Journal of Financial Economics*, 109, pp. 566-582.

Goldstein, I., and Pauzner, A. (2005), "Demand Deposit Contracts and the Probability of Bank Runs". *Journal of Finance*, 60, pp. 1293-1328.

Gromb, D., and Vayanos, D. (2002), "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs". *Journal of Financial Economics*, 66, pp. 361-407.

Hayek, F. (1945), "The Use of Knowledge in Society". *American Economic Review*, 35, pp. 519-530.

Hamilton, J. D. (1994), "Time Series Analysis". Princeton University Press.

Hellwig, C., Mukherji, A., and Tsyvinski, A. (2006), "Self-Fulfilling Currency Crises: The Role of Interest Rates". *American Economic Review*, 96, pp. 1769-1787.

Huang, J., Wei, K. D., and Yan, H. (2007), "Participation Costs and the Sensitivity of Fund Flows to Past Performance". *Journal of Finance*, 62, pp. 1273-1311.

Johnson, J. P., and Myatt, D. P. (2006), "On the Simple Economics of Advertising, Marketing, and Product Design". *American Economic Review*, 96(3), pp. 756-784.

Kabir, M. H., and Hassan, M. K. (2005), "The Near-Collapse of LTCM, US Financial Stock Returns, and the Fed". *Journal of Banking and Finance*, 29, pp. 441-460.

Kiyotaki, N., and Moore, J. (1997), "Credit Cycles". *Journal of Political Economy*, 105(2), pp. 211-248.

Kondor, P. (2012), "The More We Know about the Fundamental, the Less We Agree on the Price". *Review of Economic Studies*, 79, pp. 1175-1207.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000), "Continuous Multivariate Distributions: Models and Applications – Vol. 1". Second Edition. Wiley & Sons.

Lee, P. L. (2018), "A Retrospective on the Demise of Long-Term Capital Management". *The Columbia Law School's Blue Sky Blog*, September 10.

Lewis, T., and Sappington, D. (1994), "Supplying Information to Facilitate Price Discrimination". *International Economic Review*, 35(2), pp. 309-327.

Liu, X., and Mello, A. S. (2011), "The Fragile Capital Structure of Hedge Funds and the Limits to Arbitrage". *Journal of Financial Economics*, 102, pp. 491-506.

Matta, R., and Perotti, E. (2017), "Insecure Debt and Liquidity Runs". Unpublished.

Morris, S., and Shin, H. S. (2000), "Rethinking Multiple Equilibria in Macroeconomic Modeling". In *NBER Macroeconomics Annual*, ed. by Bernanke, B. S. and Rogoff, K., MIT Press, pp. 139-161.

Morris, S., and Shin, H. S. (2003), "Global Games: Theory and Applications". In *Advances in Economics and Econometrics: Theory and Applications*, ed. by Dewatripont, M., Hansen, L. P. and Turnovsky, S. J., Cambridge University Press.

Morris, S., and Shin, H. S. (2004), "Coordination Risk and the Price of Debt". *European Economic Review*, 48, pp. 133-153.

Morris, S., and Shin, H. S. (2016), "Illiquidity Component of Credit Risk". *International Economic Review*, 57(4), pp. 1135-1148.

Postlewaite, A., and Vives, X. (1987), "Bank Runs as an Equilibrium Phenomenon". *Journal of Political Economy*, 95(3), pp. 485-491.

Rochet, J. C., and Vives, X. (2004), "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?". *Journal of the European Economic Association*, 2, pp. 1116-1147.

Shleifer, A., and Vishny, R. W. (1997), "The Limits of Arbitrage". *Journal of Finance*, 52(1), pp. 35-55.

Shleifer, A., and Vishny, R. W. (2011), "Fire sales in Finance and Macroeconomics". *Journal* *of Economic Perspectives*, 25(1), pp. 29-48.

Spiegel, M., and Zhang, H. (2013), "Mutual Funds Risk and Market Share-Adjusted Fund Flows". *Journal of Financial Economics*, 108(2), pp. 506-528.

Stein, J. C. (2005), "Why Are Most Fund Open-End? Competition and the Limits to Arbitrage". *Quarterly Journal of Economics*, 120, pp. 247-272.

Tarashev, N. A. (2007), "Speculative Attacks and the Information Role of the Interest Rate". *Journal of the European Economic Association*, 5(1), pp. 1-36.

Vives, X. (2014), "Strategic Complementarity, Fragility, and Regulation". *Review of Financial Studies*, 27, pp. 3547-3592.