A macroeconomic model of financial instability

Fernando Mendo*

I. Introduction

Understanding the interaction between the distress of financial institutions (FI) and the real economy became of paramount importance after the last financial crisis. The literature provides two complementary approaches to guide our thinking. The first emphasizes how weak balance sheets impair FI's ability to intermediate funds due to financial frictions, which limit their capacity to expand leverage or share risks. This creates amplification and persistence. During economic downturns, asset prices decrease reducing FI's net worth, i.e. limiting their intermediation. Then, the larger misallocation of resources exacerbates the economic contraction and the drop of asset prices. Furthermore, FI's balance sheets recover slowly in the aftermath. The second approach highlights that FI are potentially exposed to runs of creditors due to the liquidity mismatch of their balance sheets: partially illiquid assets financed by short term liabilities. A run forces FI to liquidate their assets, which could lead, for example, to the inefficient termination of ongoing projects of non-financial firms.

Both phenomena played an important role during the last financial crisis. Moreover, there was a clear link between them. Losses associated with the subprime mortgages weakened balance sheets of FI which undermined creditors' confidence. First, creditors started demanding increasing interest spreads and haircuts to rollover short-term funding, and later, following the failure of Lehman Brothers, they froze their lending all together. The run (mainly on shadow banks) induced fire sales that further deteriorated FI's balance sheets and ultimately contributed to the long lasting contraction of real activity.

The goal of this project is to study financial stability in a macroeconomic model that features both a financial accelerator mechanism and runs of creditors. The focus on financial vulnerabilities in such a framework implies thinking not only on the amplification that financial constraints may generate but also on the economy's exposure to runs. This more comprehensive approach to financial stability opens a set of unexplored questions. What features characterize the situations in which the economy is exposed to runs? Do these situations coincide with the ones that feature the largest amplification due to financial constrains? How often does the economy visit those situations? How does fundamentals, e.g. volatility of productivity disturbances, influence this frequency? Which policies are effective to limit the economy's exposure to runs, e.g. safe asset provision? What determines the depth of the economic contraction after a run? This project aims to make a step forward towards answering these questions.

Building a framework that allows performing such an analysis is an important challenge. While the financial accelerator mechanism is largely developed within macroeconomic models, bank run are usually studied using partial equilibrium short-term horizon models. Moreover, most bank run models evaluate non-fundamental runs, i.e. pure coordination phenomena. These features prevent the kind of analysis we are interested on, so a first step is to embed runs into a general equilibrium framework. Recently, Gertler and Kiyotaki [2015] and Gertler et al. [2017] made progress in this direction. However, their focus on contrasting propagation mechanisms with and without runs allowed them to examine perfect-foresight solutions to a single aggregate productivity perturbation. Unfortunately, this solution strategy is not enough to analyze how often the economy is exposed to runs or which agents' decisions

*fjlopez@princeton.edu, Princeton University

lead to such exposure in equilibrium. We need the global solution for an economy constantly exposed to disturbances.

This project overcomes this challenge by setting the model in continuous time following the macro-finance literature started by Brunnermeier and Sannikov [2014] and He and Krishnamurthy [2013]. Technically, we need to take two steps forward with respect to the solution methods developed by the literature. First, systemic runs introduce aggregate jump risk which, to the best my knowledge, is a feature that no model in this literature includes. There is a simple reason for this. Aggregate jump risk breaks the key feature that makes continuous time models tractable: local perturbations. Second, the occasional presence of fire sales and the aim to explore the different situations that create vulnerability imply another technical challenge. In a nutshell, the model solutions, i.e. prices and allocations as functions of the multi-dimensional state, will feature kinks. The literature has not explored solutions with kinks beyond the single state case yet.[1]

II. Preliminary framework

First, I briefly describe a preliminary model for this project, which is a version of the two-agent-type stochastic growth model with financial frictions (constraint on equity issuances) augmented to consider a debt rollover decision. Then, I discuss some insights of this simple setup: (i) the economy's exposure to runs depends on experts' leverage and on the price-to-liquidation-value ratio of capital, (ii) the economy is vulnerable to runs only for intermediate values of bankers' capitalization, and (iii) the introduction of runs implies that the economy visit more frequently the fire sale region, i.e. where productivity shocks are amplified due to financial constraints.

Next, I explore comparative statics across economies, which uncover some striking results: (i) economies with lower fundamental volatility are more exposed to runs of creditors and amplification,[2] furthermore they present on average larger capital misallocation and lower growth, (ii) economies with a more efficient productive agent are also more exposed to runs and amplification, moreover they exhibit larger capital misallocation and larger growth volatility (and potentially lower growth). It is important to highlight that these apparently unintuitive results are not present in the version of the model that neglects to consider the possibility of runs.

a) Model.

There is a continuum of each type of agents: households (HH) and experts (or bankers). Types only differ in the efficiency of their production technologies. There are two goods: physical capital and final consumption goods. Preferences are CRRA, and there is a linear production technology for each agent: ak_t for experts, and $\hat{a}k_t$ for HH, with $a > \hat{a}$. There is a common investment technology and capital is exposed to a common aggregate productivity (or quality) shock, so individual capital evolve as

$$dk_t = (\Phi(\iota_t) - \delta)k_t dt + \sigma k_t dZ_t$$

where $\Phi(.)$ is concave and captures adjustment costs to investment per unit of capital ι_t. Note that this is the evolution of individual capital without considering capital sales or purchases, which in the case of runs will introduce a discrete change in capital holdings.

Financial markets. Agents are able to trade deposit contracts, which offer a interest rate r_t absent default. In case of default, the creditors overtake the borrowers' assets. The financial friction is an equity issuance constraint, which prevents agents from sharing the risk involved in their production decisions. There is a market for physical capital which is traded at price q_t.

Turnover. Agents face an uninsurable idiosyncratic Poisson death shock. The wealth of dying agents is equally redistributed among newborns. A fixed fraction of newborns are experts. This is to prevent experts to overtake the economy and to allow the banking sector to rebuild after a run.

[1] There is an on going project (not available yet) by Lars Hansen, Paymon Khorrami, and Fabrice Tourre at Chicago University which is also dealing with kinks in an economy with multiple states.

[2] That is, the economy spends more time in the region where productivity shocks are amplified.

Run of creditors. At the beginning of each "dt-period," households decide whether to rollover their deposits or not. A run can occur when liquidation value of assets is less than the value of deposits. In such cases, there is equilibrium multiplicity. In case a run is realized, all bankers' net worth is erased and the banking sector rebuilds itself due to newborn bankers.

This version of the model uses the following two assumptions for simplicity.

1. *Sunspots.* Whenever a run is possible, it is coordinated by a sunspot, which arrives with a Poisson intensity that is an increasing function of the percentage loss depositors face if the run is realized.
2. *Collective failure.* An individual banker cannot survive if all others bankers fail.

The first assumption is a simple way to resolve the multiplicity inherent to runs. The second is necessary to avoid bankers stop using leverage as soon as there is a positive probability of a run. This would happen because death (or zero consumption) is infinitely costly for agents. In a version where bankers survive runs, e.g. becoming households with a fraction of their assets, the second assumption can be relaxed.

b) Solution.

I omit the equilibrium characterization due to space considerations. Instead, I briefly describe how the exposure to runs influence agents' decisions. I also discuss the features of the equilibrium without runs that are necessary to understand the financial stability analysis that follows.

Agents' decisions. When runs are possible, experts act more impatiently since they internalize their higher death probability. They increase their consumption rate but their (myopic) demand for capital is not affected. Meanwhile, households reduce their demand for capital because losses from capital holding are larger than those from deposits once a run is realized. The reason behind is that deposits are partially protected by borrowers' net worth.

Equilibrium without runs. This economy can be reduced to a single state: the experts' share of wealth η. The equilibrium features two regions. For high values of η, experts decide to absorb all capital in the economy and they are compensated for their exposure to productivity risk with a positive risk premium. In this region, the capital price is fairly constant. For low values of η, the risk compensation that experts would require to hold all capital is so large that it is optimal for the unproductive agent to hold some of it. In this fire-sale region, capital price is increasing with η because larger wealth for experts reduces their effective risk aversion, which increases their capital demand.

c) Financial stability analysis.

Equilibrium exposure to runs. The loss that households face on their deposits in case of a run depends on 2 factors: the liquidation value of the asset, and banks' balance sheet. In particular, we can write it as

$$\text{\% loss of deposits} = \max\left\{0, 1 - \left(\frac{\text{Liquidation value}}{\text{Capital price}}\right)\left(\frac{\text{Experts' Leverage}}{\text{Experts' Leverage-1}}\right)\right\}$$

which is an increasing function of capital's price-to-liquidation-value ratio and banks' leverage.

As consequence, the economy is exposed to runs only for intermediate levels of banks' net worth share. When η is too low the potential drop on capital price (to its liquidation value) is too small to erase banks' net worth. When η is too big bankers operate with little leverage so even a large drop in capital price would not translate in losses to depositors. The left panel of Figure 1 presents the potential loss as function of η.

When a bank run is possible, the s*unspots assumption* implies that the economy's exposure to runs (i.e. Poisson intensity of sunspots) is single peaked as function of bankers' wealth share. Further, the economy's largest exposure coincides with the threshold wealth share level that allows banks to hold all capital. The intuition behind is illustrated in the right panel of Figure 1. In the fire-sale region, the economy's exposure to runs is driven by capital's price-to-liquidation-value ratio[3] (increasing in η), while in the non-fire-sale region, it is driven by leverage (decreasing in η).

Amplification of productivity shocks and runs. The amplification of productivity shocks is summarize by the sensitivity of capital price to those shocks, which the literature refers to as endogenous volatility σ_q. In fact, it is the endogenous part of the total capital return volatility, $\sigma + \sigma_q$. This endogenous volatility is higher in the fire-sale region because of the balance sheet channel: positive shocks translate into larger net worth for experts, which implies a more efficient capital allocation, and therefore a higher capital price.

The possibility of runs makes the economy more exposed to amplification in two ways. First, the economy spends a larger fraction of time in the fire-sale region, i.e. the region where bankers are poorly capitalized. The reason is that the economy does experience runs in equilibrium. Runs erase the banking sector's net worth completely, and its recovery is slow and volatile. The left panel of Figure 2 shows the change in the stationary distribution due to the introduction of runs. The larger mass in the fire-sale region when considering runs is evident.

Second, whenever households hold some capital, the exposure to runs increases endogenous volatility. The reason is the explained reduction in capital demand by households. In equilibrium, this implies larger leverage for bankers, which increases the sensitivity of its wealth share to productivity shocks. This reinforces the financial accelerator mechanism increasing the amplification of the productivity shocks (endogenous volatility). The right panel of Figure 2 shows the increase in endogenous volatility for region with fire-sale that is exposed to runs (between the green and red vertical lines).

d) Comparative statics across economies.

Low volatility of productivity shocks leads to financial fragility. As exogenous volatility σ decreases, the economy's exposure to runs and the average amplification effects increase. Moreover, the economy presents larger average capital misallocation and lower growth rate.

The intuition behind is the following. As σ declines, bankers' capital demand raises which in equilibrium translates into larger capital price and banks' leverage. These two forces increase the economy's exposure to a systemic run, which reduces households' capital demand. The latter reinforces the increase of bankers' leverage leading to even further exposure to runs. The more frequent runs generate that bankers spend more time with little net worth unwilling to hold all capital (fire sale region). Therefore, the economy spends large spans of time with substantial endogenous volatility, large misallocation, and low growth.

Figure 3 illustrates the effects of decreasing σ. Top panels show the increasing financial fragility: larger exposure to runs (left panel), and more time spend in fire sale region (right panel, see height of highlighted point which indicates the threshold for fire-sales). The middle panels present the reason behind the larger exposure: larger leverage (left) and higher price-to-liquidation value (right). Bottom panels show how capital misallocation changes in the model without runs (left) and with them (right). The panels show the physical capital share of bankers φ as a function of the CDF of the stationary distribution of η. Then, the average φ is the area below the curve. The left panel shows a slight improvement of capital allocation as σ decreases for the economy without bank runs, while the right panel shows a strong opposite effect when considering runs.

Large bankers productivity leads to financial fragility. As the productivity of bankers a increases, the economy's exposure to runs and the average amplification effects increase. Further, the economy presents larger average capital misallocation and effects over growth are mixed.

As a increases, bankers' demand for capital increases and the adjustment is mainly through an increase in capital price.[4] The larger capital price-to-liquidation-value ratio raises the economy's exposure to runs. More frequent runs translate into more time spent in the fire sale region. This increases average

[3]In this simple economy, the liquidation value of capital is constant an equal to the capital value when households own all wealth (and therefore all capital).

[4]Bankers' leverage is less sensitive due to an opposing force: larger a increases the goods supply so equilibrium demands either larger capital price or lower capital share for bankers, i.e. less leverage.

endogenous volatility, and capital misallocation. The effect over growth is mixed since the economy spends more time in the fire sale region where growth is low but when it does escape from this region the larger a implies a higher capital price, investment and growth.

Figure 4 illustrates the effects of increasing a. Top panels show the increase in financial fragility: larger exposure to runs (left panel), and more time spend in fire sale region (right panel). Middle right panel presents the reason behind the larger exposure: increase in capital price to liquidation value. Bottom panels show how bankers' capital share is almost constant for the case without runs (left), but sharply decreases once runs are allowed (right).[5]

III. Discussion

I have presented a simple macroeconomic model that incorporates both a financial accelerator mechanism and runs of creditors, and yet is tractable enough to be solved globally. I used this model to explore two dimensions of financial stability: the economy's exposure to runs and the endogenous amplification of productivity shocks. The analysis helped us to understand in which situations the economy is exposed to runs and how the possibility of runs has consequences for the balance sheet amplification mechanism. Moreover, the model uncovers how characteristics that may seem positive for the economy, such as low fundamental volatility and an efficient experts, can lead to weaker outcomes because they generate a large exposure to systemic runs.

There are several immediate areas for further work. First, incorporating to the analysis a safe asset, i.e. an asset not exposed to runs or productivity shocks, would allow exploring which are the consequences of policies, such as changes in the supply of government bonds, over financial stability. There is a potential tension on the impact of safe assets over welfare. It provides an insurance mechanism against runs, but it may also drive up the cost of funding of financial institutions precipitating a run.

Evidently, the interest in examining government policies requires a formal welfare analysis. This is a second direction for future progress. Models with an endogenous exposure to a death shock, e.g. runs for bankers, are unsuitable for welfare comparisons. The reason is that the time varying exposure to death introduces distortions through the arbitrary value assign to agents' utility flows after he exits the economy (usually zero). A version of the model in which bankers survive runs avoids this issue.

Also, I limited the discussion above to runs on the entire financial sector, however, in reality this sector encompasses heterogeneous financial institutions and runs happen only on a subset of them, usually the ones with the weakest balance sheets, e.g. shadow banks. This is a third venue for future work. A simple approach to capture the heterogeneous exposure of financial institutions is to introduce idiosyncratic productivity jump risk. Then, we can explore the situations in which the economy is exposed to runs of different depths (according to the share of failing institutions).

Finally, given the insights of the comparative static exercises with respect to bankers' productivity and exogenous volatility, a natural next step is to formalize their evolution behavior with an adequate stochastic process. Then, we can explore how financial fragilities evolve as the economy's fundamentals change. Beyond characterizing the regions with exposure to runs, it would be of particular interest to capture the behavior observed just before the run on the shadow banking system materialized, i.e. increasing spreads and declining net worth of financial institutions.

This discussion emphasized the venues for progress to enhance our understanding of financial fragility and the potential policies to avoid it. A transversal line of future work is to discipline the numerical exercises using, for example, a simulated methods of moments procedure. It is also worth to mention that some of the discussed venues involve an important technical challenge.

[5]See discussion of Figure 3 for detail on how to read bottom panels.

References

Markus K. Brunnermeier and Yuliy Sannikov. A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421, 2014. doi: 10.1257/aer.104.2.379. URL `http://dx.doi.org/10.1257/aer.104.2.379`.

Mark Gertler and Nobuhiro Kiyotaki. Banking, Liquidity and Bank Runs in an Infinite Horizon Economy. *American Economic Review*, 2015.

Mark Gertler, Nobuhiro Kiyotaki, and Andrea Prestipino. A Macroeconomic Model with Financial Panics. 2017.

Zhiguo He and Arvind Krishnamurthy. Intermediary asset pricing. *American Economic Review*, 103 (2):732–770, 2013. ISSN 00028282. doi: 10.1257/aer.103.2.732.

Figure 1: Equilibrium exposure to runs

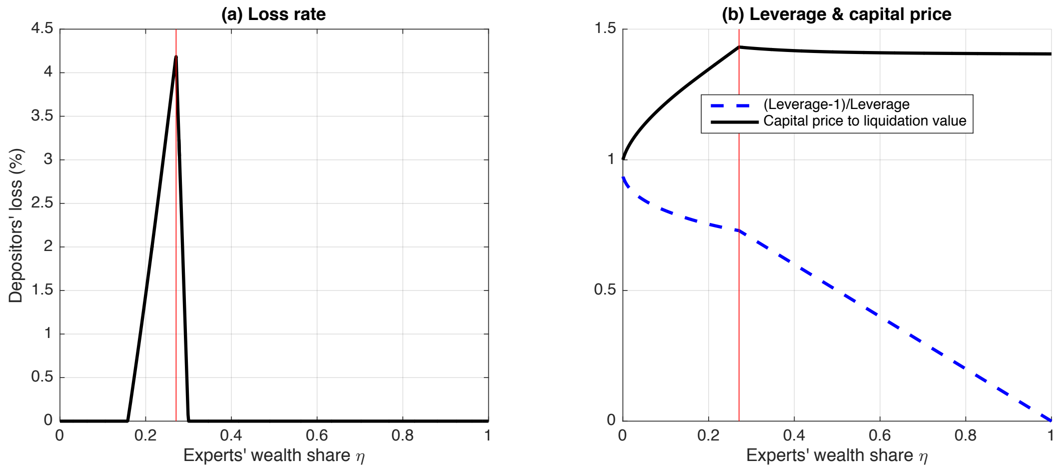

Figure 1. Panel (a) shows the percentage loss that depositors face in case the run is realized. Panel (b) presents the factors that influence this potential loss: leverage (monotone transformation) and the capital price-to-liquidation-value ratio. The vertical red line separates the fire-sale region (low η) from the region where bankers hold all capital (high η).

Figure 2: Amplification of productivity shocks

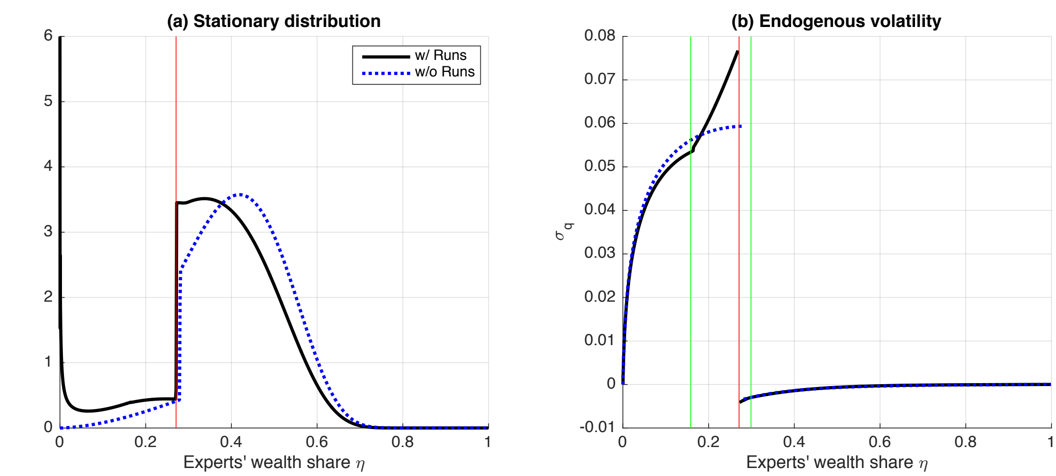

Figure 2. Panel (a) contrasts the stationary distribution with and without allowing for run. Panel (b) presents the endogenous volatility for both cases. The vertical red line separates the fire-sale region (low η) from the region where bankers hold all capital (high η). The green vertical lines limit the region where the economy is exposed to runs.

Figure 3: Exogenous volatility and financial stability

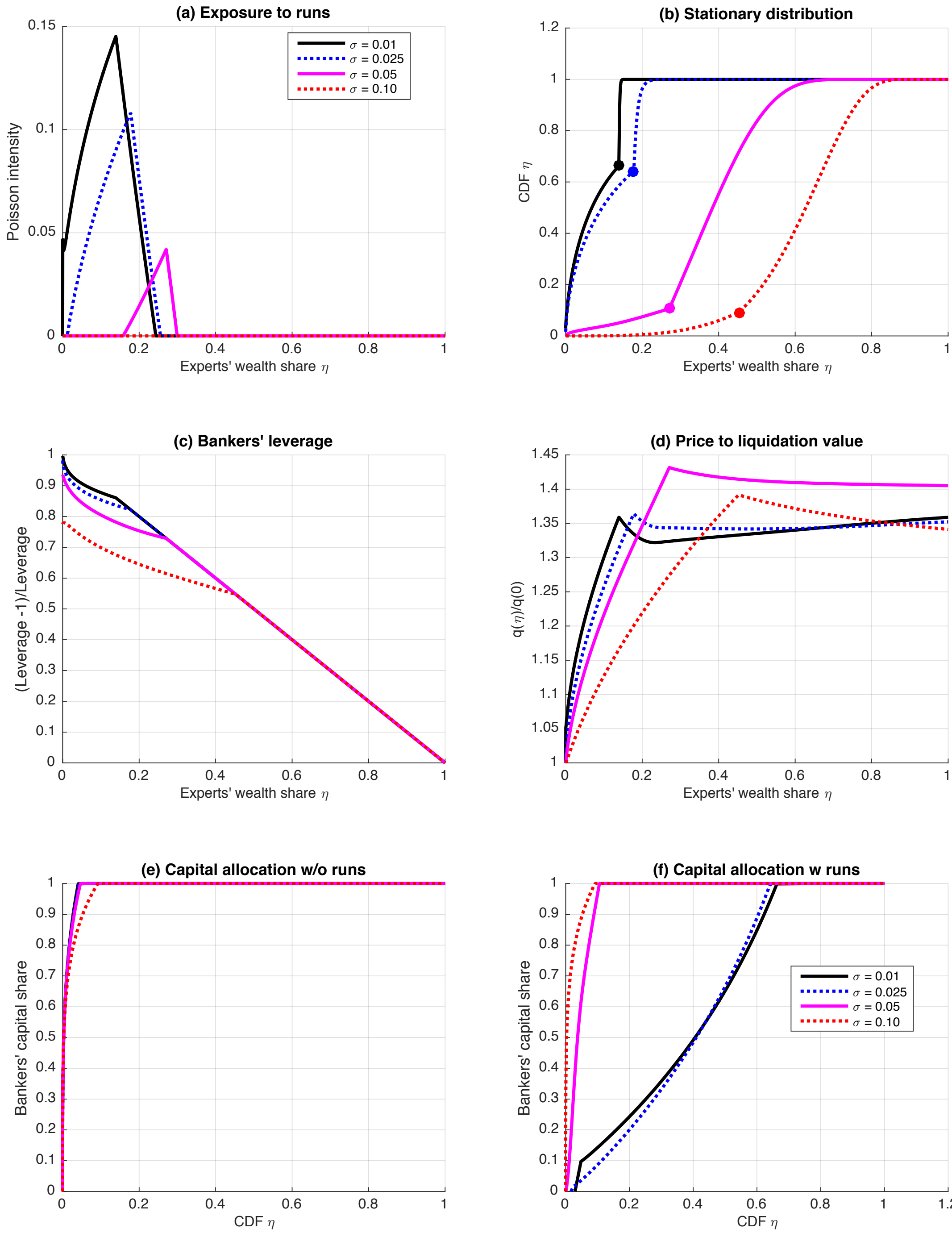

Figure 3. Panel (b) presents the CDF of stationary distribution and the point corresponding to the limit between the fire-sale and the non fire-sale region is highlighted. The height of this point in the graph indicates the share of time spend in the fire sale region. Bottom panels (e) and (f) use the CDF of η in the x-axis. Then, the average of the variable shown is the area below the curve.

Figure 4: Bankers' productivity and financial stability

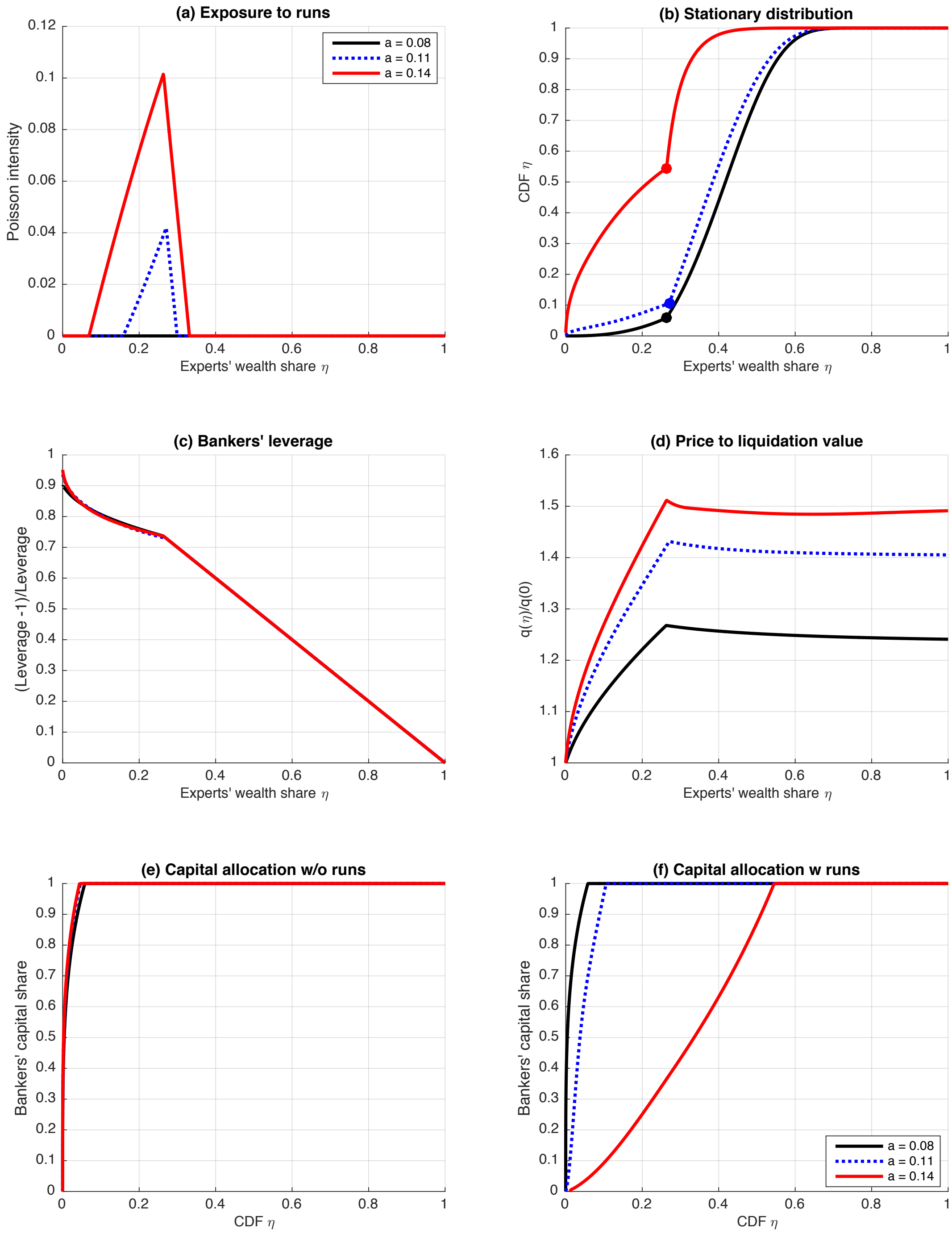

Figure 3. Panel (b) presents the CDF of stationary distribution and the point corresponding to the limit between the fire-sale and the non fire-sale region is highlighted. The height of this point in the graph indicates the share of time spend in the fire sale region. Bottom panels (e) and (f) use the CDF of η in the x-axis. Then, the average of the variable shown is the area below the curve.