THE UNIVERSITY OF CHICAGO

ESSAYS ON FINANCIAL INTERMEDIARIES AND MONETARY POLICY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

AND

THE FACULTY OF THE UNIVERSITY OF CHICAGO

BOOTH SCHOOL OF BUSINESS

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

STEFANO PEGORARO

CHICAGO, ILLINOIS

JUNE 2020

Dedicata a mia madre Augusta, mio padre Giuliano e mio fratello Alessandro,

senza i quali non avrei mai potuto raggiunto questo traguardo.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

ABSTRACT

Chapter 1. Previous literature documents that mutual funds' flows increase more than linearly with realized performance. I show this convex flow-performance relationship is consistent with a dynamic contracting model in which investors learn about the fund manager's skill. My model predicts that flows become more sensitive to current performance after a history of good past performance. It also predicts that the effect of past performance on the current flow-performance relationship is weaker for managers with longer tenure. I consider an optimal incentive contract for money managers, and I provide an explanation for common compensation practices in the industry, such as convex pay-for-performance schemes and deferred compensation. In the optimal contract, flows become more sensitive to performance when the manager faces stronger incentives from the compensation contract. With learning, the manager's incentives become stronger after good performance, so that a manager exerts more effort when his assessed skill is higher. However, the relation between past performance and incentives becomes weaker over the manager's tenure. Using mutual fund data, I test the predictions of the model on the dynamic behavior of the flow-performance relationship, and I find empirical support for the theory.

Chapter 2. We test through which channels quantitative easing affects the prices and issuance of securities. We exploit the announcement of the corporate bond purchase program by the European Central Bank, and we study the impact of the announcement on the cross section of European corporate bonds. We find that as the Central Bank increased the demand for bonds eligible for the program, eligible firms responded by substituting the issuance of ineligible bonds with the issuance of

eligible bonds. As a result, bond prices were unaffected by their eligibility status, and all firms increased total issuance to the same extent. We show that monetary policy affected bond prices through a risk channel. Prices increased significantly more for bonds and firms that were exposed to higher levels of risk and uncertainty. However, risky firms did not issue more in response.

Chapter 3. I develop a continuous-time game between a population of investors and an intermediary whose type is private information and whose portfolio allocation is neither observable nor contractible. I define and characterize a sequential equilibrium of the game and solve for a Markovian equilibrium where investors' posterior beliefs are the key state variable. In my model, demand for riskless assets undergoes dramatic changes that resemble the episodes of flight to safety observed during financial crises. I show that a risk-neutral intermediary chooses a portfolio that minimizes risk when beliefs are near the threshold below which the intermediary is terminated.

CHAPTER 1

FLOWS AND PERFORMANCE WITH OPTIMAL

MONEY MANAGEMENT CONTRACTS

1.1 Introduction

Mutual funds face a double challenge. On the one hand, they cope with volatile flows of money as investors react to funds' performance: Well-performing funds experience money inflows, whereas poorly-performing funds experience money outflows. On the other hand, funds may struggle to generate good performance if their portfolio managers face inadequate incentive schemes. These two challenges are tightly connected. Funds collect fees on their assets under management. They therefore aim to maximize performance in order to increase money inflows and fee revenues. However, funds need to delegate investment decisions to managers whose objective is not to maximize performance but to pursue their own interests. In this paper, I study the connection between these two challenges and show how funds can overcome them using optimal money management contracts.

I develop a dynamic contracting model that explains common patterns in the money management industry, including a convex relation between fund flows and performance, as well as convex pay-for-performance schemes for portfolio managers. Existing literature studies these patterns in isolation, often viewing them as puzzling, and sometimes viewing them as problematic. Brown et al. (1996) and Chevalier and Ellison (1997) interpret the convex flow-performance relationship as an implicit incentive scheme for mutual funds, and investigate its implications for funds' risk-

taking behavior. Basak and Pavlova (2013), Cuoco and Kaniel (2011), and Panageas and Westerfield (2009) study how common compensation schemes distort managers' portfolio choices. Using a mechanism-design approach, I show that a convex flow-performance relationship and a convex compensation scheme are jointly consistent with an optimal dynamic contract for money managers. To provide empirical support for my theory, I derive novel statistical predictions on the dynamic behavior of the flow-performance relationship and test them in mutual fund data.

In the model, I consider the two layers of incentives that characterize the mutual fund industry. In the first layer, competitive investors supply capital and pay proportional fees to a fund advisor who represents the fund family or the fund management company. In the second layer, the advisor hires a portfolio manager and sets the terms of the manager's compensation contract. Because the advisor captures the value added of the fund through fee revenues, she faces implicit incentives to maximize performance and assets under management. To align the manager's incentives to her owns, she offers an explicit performance-based compensation contract to the manager.

To generate an increasing and convex relation between fund flows and performance, the model relies on two assumptions. First, the manager possesses an unknown skill to generate excess returns. The manager, the advisor, and the investors learn about the manager's skill by observing realized returns. According to this assumption, investors expect better future performance from a manager who performed better in the past. Second, the manager is subject to moral hazard in his private choice of costly effort: He may exert low effort and reduce returns for investors.

The advisor designs an optimal incentive contract to prevent this possibility. In the optimal contract, the advisor specifies the incentives of the manager as an increasing function of the manager's expected performance, so that a manager exerts more effort when his assessed skill is higher. Moreover, flows into and out of the fund become more sensitive to performance when the manager faces stronger incentives. Fund flows thus become more sensitive to current performance when future expected performance increases. Therefore, as good returns accumulate and investors expect increasingly better future performance, fund flows become increasingly more sensitive to current performance. As a result, over any period of time, cumulative fund flows respond in a positive and convex way to cumulative performance. The empirical literature has repeatedly documented a positive and convex response of fund flows to performance (Chevalier and Ellison, 1997; Del Guercio and Tkac, 2002; Sirri and Tufano, 1998).

The model produces a managerial compensation scheme that reflects three common practices in the money management industry. First, managers are compensated for their performance. Second, similar to capital flows, the manager's compensation becomes more sensitive to performance after a history of good returns, thus resulting in a convex compensation scheme on an annual basis. These first two results are consistent with the widespread use of convex pay-for-performance contracts in the industry (BIS, 2003; Ma et al., 2019). According to the model, advisors opt for these contracts so that managers with higher assessed skill face stronger incentives to exert effort. Third, the optimal contract includes a deferred compensation feature. After good performance, the advisor partially postpones the delivery of the

promised compensation in order to provide stronger incentives to the manager. In the money management industry, several practices effectively postpone the payout of performance-based compensation to future periods.[1]

I provide empirical support for my theory by testing a novel prediction of the model regarding the dynamic behavior of the flow-performance relationship: Fund flows react more strongly to current performance after a history of good performance. I measure a mutual fund's performance by comparing its return with the average return of funds with the same investment objective. For every month, I compute the average past performance over the previous months. At a monthly frequency, past performance has a positive and statistically significant effect on the extent to which fund flows respond to current performance.

I test an additional novel prediction of the model, which establishes a connection between the dynamic behavior of the flow-performance relationship and the tenure of the manager. Under the assumption that the advisor and the manager can fully commit to the terms of the optimal long-term contract, the slope of the flow-performance relationship depends more weakly on past performance if the manager has a longer tenure. I verify this prediction holds in mutual fund data. According to the model, the ex ante optimal contract imposes constraints on the incentives and on the skill of the manager as the manager's tenure becomes longer. If the manager's skill plays a minor role in generating returns, past performance provides less information about future performance, thus weakening the link between past performance and the slope

1. Ma et al. (2019) document that 30% of the mutual fund managers in their sample are subject to explicit deferred compensation contracts. Moreover, they find that managers' bonuses depend on their average performance over multiple years in the past. This last practice effectively implements a deferred compensation scheme.

of the flow-performance relationship.[2]

I derive my results in a continuous-time contracting problem with learning about the manager's skill. This problem poses some challenges, which I overcome by using duality methods. Given an optimal incentive contract with learning, the manager could acquire an ex post information rent after shirking. Intuitively, a manager who shirks and appears unskilled has better career prospects than a manager who is actually unskilled. If a contract induces larger information rents, the manager has weaker incentives to maximize returns and the advisor obtains lower revenues. Therefore, one could formulate an optimization problem in which the advisor controls the manager's information rent as an independent state variable. Unfortunately, this problem cannot be feasibly solved. To overcome this challenge, I use duality methods and offer a tractable and intuitive formulation of the contract-design problem. In the optimal contract, the advisor commits to ex post inefficient incentives in order to reduce the ex ante information rent of the manager. In the dual formulation, this commitment is captured by a multiplier that, over time, distorts the terms of the contract towards a lower risk exposure for the manager. By considering the dynamics of the multiplier and of beliefs, I provide intuitive interpretations for my results on the flow-performance relationship and on the manager's compensation scheme.

The rest of the paper is organized as follows. In section 1.2, I review the related literature. In section 1.3, I present the setup of the model. In section 1.4, I

---

2. Almazan et al. (2004) provide evidence that more experienced managers are subject to more investment constraints. My model suggests this evidence could simply represent the outcome of an optimal contract that restricts the use of the manager's skill over time.

characterize the optimal contract. In section 1.5, I show the implications of the optimal contract for flows, performance, and compensation and provide the economic intuition behind the results. In section 1.6, I empirically test model's predictions. Section 1.7 concludes.

In Appendix 1.A, I show the model's key results hold also when I relax the assumption that the manager and the advisor fully commit to the terms of the optimal contract. Appendix 1.B contains a two-period model that illustrates the trade-off between ex ante and ex post efficiency. All proofs are in Appendix 1.C, Appendix 1.D, and Appendix 1.E. Appendix 1.F contains robustness checks for the empirical analysis.

1.2   Related Literature

My paper extends the current asset management literature by studying the connection between agency frictions, managerial compensation, and the flow-performance relationship. Existing models of asset management (Basak and Pavlova, 2013; Cuoco and Kaniel, 2011; Kaniel and Kondor, 2013; Vayanos and Woolley, 2013) focus on the implications of common fee schedules on portfolio choices and asset prices. Compared to this literature, my model builds on the recent empirical evidence in Ibert et al. (2018) and Ma et al. (2019), who show that portfolio managers face compensation contracts that differ substantially from the fee revenues that fund management companies receive. Although my paper does not study the asset-pricing implications of managerial incentive schemes, it provides a theoretical foundation for common compensation practices in the industry. My paper is therefore related to the lit-

erature that adopts a mechanism-design approach to study investment delegation (Bhattacharya and Pfleiderer, 1985; Cadenillas et al., 2007; Dybvig et al., 2010; He and Xiong, 2013; Heinkel and Stoughton, 1994; Ou-Yang, 2003; Palomino and Prat, 2003). I further develop this literature by studying the connection between managerial incentives and the flow-performance relationship.

Compared to previous theoretical studies of the convex flow-performance relationship, my research highlights the dynamic nature of such a relationship. In particular, the model produces a relation between flows and performance that depends on the history of past performance. This dynamic aspect of the flow-performance relationship is novel to the literature. My model therefore complements Berk and Green (2004), in which the flow-performance relationship changes as a function of the fund's age, but not of past performance. To derive my results, I rely on learning about the manager's skill and on the agency frictions inside the fund. I abstract from other determinants of the flow-performance relationship that previous studies have considered, for example, decreasing returns to scale (Berk and Green, 2004), changes in investment strategies (Lynch and Musto, 2003), participation costs (Huang et al., 2007), or bias in the social transmission of information (Han et al., 2018).

In my empirical analysis, I stress the dependence of the flow-performance relationship on the history of the fund's past performance and on the manager's incentives. My analysis therefore differs from Franzoni and Schmalz (2017), who study how the flow-performance relationship changes with aggregate risk factors, rather than with the fund's idiosyncratic performance. My research complements also the analysis in Chevalier and Ellison (1997). Whereas they study how the shape of the flow-

performance relationship differs across young and old funds, I study how the history dependence of the flow-performance relationship varies with the tenure of the manager. Although my empirical analysis is primarily related to the extensive literature on the flow-performance relationship (Chevalier and Ellison, 1997; Del Guercio and Tkac, 2002; Ippolito, 1992; Sirri and Tufano, 1998; Zheng, 1999), I design my empirical tests on the basis of a model that connects fund managers' incentives to the flow-performance relationship. In this sense, my paper adds a new perspective to the literature that explores the connection between managers' incentives and fund performance (Agarwal et al., 2009; Almazan et al., 2004; Chen et al., 2008; Chevalier and Ellison, 1999; Khorana et al., 2007).

From a modeling perspective, my paper builds on the literature about dynamic contracting with learning (Bergemann and Hege, 2005; DeMarzo and Sannikov, 2017; Halac et al., 2016; He et al., 2017; Hörner and Samuelson, 2013; Prat and Jovanovic, 2014). I contribute to this literature along two dimensions. First, I allow for capital flows and for smooth changes in the rate of learning. By doing so, I obtain testable predictions on the relation between capital flows, performance, and managers' tenure. Second, I show how to use duality methods to overcome technical challenges posed by this class of contracting models. In particular, I use duality methods to solve a dynamic programming problem with an endogenous bound on the state space, which would be otherwise unfeasible. Sannikov (2014) and Miao and Zhang (2015) also use duality methods to solve contracting problems, although with the purpose of obtaining linear partial differential equations.

1.3   Model Setup

I consider a setting that involves a portfolio manager, a fund advisor, and a population of investors as players. No player knows the true alpha-generating skill of the manager, and a standard agency friction exists within the fund, because the manager could shirk and gain a private benefit at the expense of investors. Time is continuous and starts at 0.

1.3.1   Players

I consider the following players: a population of investors, one fund advisor (the principal/she), and a portfolio manager (the agent/he). I use the term "fund" to refer to the organization formed by the portfolio manager and the fund advisor. The advisor collects capital from investors on the spot market and hires the manager to actively manage this capital.

Investors are risk neutral, competitive, and cannot commit to long-term contracts. They interact with the advisor through a series of spot contracts that, at every time t, specify the assets investors supply to the fund, K_t, and the proportional fee the advisor receives, f_t. Investors then collect the returns that the fund produces, R_t. Through their interaction with the fund, investors obtain a utility

$$\mathrm{E}\left[\int_0^\infty e^{-rt}\left(K_t dR_t - (f_t + r)K_t dt\right)\bigg|\mathcal{F}_0\right],$$

where $r > 0$ is the market risk-free rate and represents both the discount rate of

9

investors, as well as their opportunity cost of capital.[3] Investors are therefore willing to provide any amount of capital as long as they expect the net-of-fee return to weakly exceed the risk-free rate. $\mathcal{F}_0$ captures the initial information, which is common across all players.

Through most of the paper, I assume that the fee is variable over time and that K_t represents the amount of assets that the fund actively manages. Alternatively, one could assume, as in Berk and Green (2004), that the fee is fixed at some sufficiently low value, $\bar{f}$, and that investors supply additional capital, $\bar{K}_t$, which the fund invests in a passive benchmark. Investments in the passive benchmark can be easily monitored, so they are not subject to the moral-hazard friction that I describe shortly. These two assumptions yield equivalent outcomes, as long as the passively and actively managed parts of the portfolio are contractible and the implied transfers coincide, that is, $f_t K_t = \bar{f}(K_t + \bar{K}_t)$.

The advisor collects the fees paid by investors, $f_t K_t$, and offers a compensation $\tilde{C}_t$ to the manager. The advisor is risk neutral and her objective is to minimize the cost of running the fund,

$$\mathrm{E}\left[\int_0^\infty e^{-rt}(\tilde{C}_t - f_t K_t)dt \,\middle|\, \mathcal{F}_0\right].$$

Unlike the investors, the advisor has the power to commit to a long-term contract with the manager.[4] Investors observe this long-term contract and understand the

---

3. These preferences reflect the assumption that investors are active in a complete asset market. The fund offers an idiosyncratic return that they are able to fully diversify. The fund's returns are thus uncorrelated with their aggregate consumption and carry no risk premium for investors.

4. I relax the full-commitment assumption in Appendix 1.A.

consequences of the contract on the manager's incentives. In particular, investors adjust their supply of capital and their willingness to pay fees in response to the terms of the contract. Therefore, in designing the contract, the advisor accounts not only for the incentives of the manager, but also for the investors' responses.

The manager controls the fund's active portfolio. However, he cannot be directly monitored by the advisor. In particular, the advisor cannot verify whether the manager exerted full effort in making his investment and trading choices. Because of the imperfect monitoring, a moral-hazard friction emerges. More formally, I assume the agent may shirk at rate m_t and reduce the fund's cash flow rate by $m_t K_t$. The manager obtains a private consumption value of $\lambda m_t K_t$ from shirking, where $1 - \lambda \in (0, 1)$ represents the inefficiency of shirking. Full effort coincides with $m_t = 0$, and the consumption value of shirking can be equivalently interpreted as the cost of effort. If a fund manager consumes $(C_t)_{t \geq 0}$ and shirks at a rate $(m)_{t \geq 0}$, his expected utility is given by

$$V_0 = \mathrm{E}\left[\int_0^\infty e^{-\delta t} u(C_t + m_t \lambda K_t)\, dt \,\middle|\, \mathcal{F}_0\right], \tag{1.1}$$

where $(K_t)_{t \geq 0}$ are the fund's assets under management. I assume $\delta \geq r$ and $u(x) = \frac{x^{1-\rho}}{1-\rho}$ with $\rho \in (0, 1/2)$.[5]

---

5. The assumption that $\rho < 1/2$ is needed to obtain a finite solution to the model. If $\rho > 1/2$, the manager's marginal utility of consumption would decline quickly enough that the advisor would find it profitable to give infinite capital and infinite consumption to the manager and overcome the incentive problem. This assumption can be relaxed if I extend the model to allow the manager to privately save. However, this extension would substantially complicate the model along several dimensions. Moreover, it would not add any additional insight about the economic mechanism that determines the flow-performance relationship.

The manager possesses a specific skill, which is unobservable to the advisor and the investors, as well as to the manager himself. However, he can generate informative signals about his skill through costly experimentation. For example, he can search for profitable investment opportunities or implement new trading strategies. If the manager is skilled, he will produce additional excess returns. If he is not, his experimentation efforts will be worthless. By observing realized returns, all players will be able to learn over time whether the manager possesses a superior investment skill. I assume experimentation can be represented by a variable, η_t, that takes values in a bounded set, $\eta_t \in [0, \bar{\eta}]$. Experimentation is fully observable and contractible. To introduce costs for experimentation, I assume experimentation reduces the consumption value of the manager's compensation. If the manager receives compensation $\tilde{C}_t$ from the advisor, but he is required to undertake experimentation η_t, his final consumption is given by

$$C_t = q(\eta_t)\tilde{C}_t, \tag{1.2}$$

for a function $q(\cdot)$ such that $q(\cdot) \in (0, 1)$, $q'(\cdot) < 0$ and $q(\bar{\eta}) \geq \lambda$. We can interpret the quantity $(1 - q(\eta_t))\tilde{C}_t$ as the cost of searching for new investment ideas. If the manager searches more assiduously for new investment ideas, then his experimentation rate η increases, and the consumption value of his compensation decreases.

1.3.2 Returns

The advisor hires the manager to actively manage the fund's portfolio and generate excess returns. However, the manager's skill is unknown to all players, including the manager. I assume the manager could be either skilled or unskilled, as indexed by his hidden type $h \in \{0, 1\}$. If the manager experiments with investment ideas and trading strategies, his skill will be reflected in the fund's returns. Players then learn over time about the manager's skill by observing the fund's performance.

I assume returns follow a stochastic process,

$$
dR = (r + \mu(\eta_t, h, m_t))dt + \sigma dW_t
$$
$$
\mu(\eta_t, h, m_t) = \alpha + \sigma \eta_t h - m_t,
$$

(1.3)

where r is the risk-free rate, and $\alpha \geq 0$ and $\sigma \geq 0$ are known parameters.

Returns depend on the uncertain skill of the manager, h, and on the manager's hidden action, $m_t \geq 0$, which is positive if the agent shirks and does not exert full effort in managing the assets. If a manager is skilled ($h = 1$), he obtains superior returns by experimenting, that is, by setting $\eta_t > 0$. If he is unskilled ($h = 0$), his experimentation efforts will not be reflected in returns. By shirking at rate m_t, the manager reduces cash flow for investors by $m_t K_t$. However, his private benefit of shirking is only $\lambda m_t K_t$ for $\lambda < 1$. Shirking is therefore inefficient because the manager destroys more value that what he obtains. The advisor, in order to maximize her revenues, designs a contract that enforces full effort. When investors observe this contract, they understand that the manager has incentives to exert full effort and they account for these incentives when deciding how much capital to provide to the

13

advisor for a given level of fees.

Besides affecting the distribution of returns, experimentation determines the information content of returns as a signal for the manager's skill. Experimentation η_t coincides with the signal-to-noise ratio of returns at t, which is defined as

$$\frac{\mu(\eta_t, 1, m_t) - \mu(\eta_t, 0, m_t)}{\sigma}.$$

This quantity measures how informative returns are regarding the manager's skill. Suppose a skilled manager produces, on average, much larger returns than an unskilled manager, that is $\mu(\eta_t, 1, m_t) \gg \mu(\eta_t, 0, m_t)$. In this case, a good (bad) return realization will be a strong signal that the manager possesses high (low) skill. However, if the volatility of returns, σ, is very large, a skilled manager could generate a very poor return due to bad luck, whereas an unskilled manager could deliver a superior return due to good luck. Therefore, volatility reduces the information content of return signals. By increasing experimentation, η_t, the fund increases the signal-to-noise ratio of returns and hence generates more information about the manager's skill. Although experimentation is beneficial in the short run, I show that in the long run future experimentation worsens the moral-hazard problem. In the optimal contract, the advisor will therefore trade off the benefits of current experimentation with the costs of any future experimentation that she promises.

1.3.3 Contracting Environment and Learning

The two main elements of the model are the incentive contract between the advisor and the manager, and the learning process. Returns constitute the only source of information for the players. First, returns allow players to draw inference about the manager's skill. Second, because return realizations change depending on the hidden action of the manager, they can be used as the basis of an incentive contract. Therefore, the terms of the contract between the manager and the advisor, as well as the spot contracts between the advisor and the investors, can solely depend on the history of returns. In this framework, players' information is generated by the history of returns and I denote with $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by the history of returns. Let $(R_s)_{0 \leq s \leq t}$ denote the history of returns up to time t, then $\mathcal{F}_t = \{(R_s)_{0 \leq s \leq t}\}$ is the smallest σ-algebra for which $(R_s)_{0 \leq s \leq t}$ is measurable.

A contract between the advisor and the manager specifies the manager's consumption, the size of his actively managed portfolio, and the experimentation that he undertakes. Moreover, for completeness, I assume the contract also specifies the effort that the advisor expects. Although the agent's effort cannot be verified, the advisor will form a conjecture about the manager's hidden action at any point in time. I let the contract specify this conjecture. To keep the notation parsimonious, I omit compensation from the definition of contract. Given consumption C_t and experimentation η_t, the compensation $\tilde{C}_t$ of the manager is determined by equation (1.2).

Definition 1.1 (CONTRACT). *A contract $\mathcal{C}$ is a set of $\mathcal{F}_t$-adapted processes* $((C_t)_{t \geq 0}, (K_t)_{t \geq 0}, (\eta_t)_{t \geq 0}, (m_t)_{t \geq 0})$.

15

Although the advisor cannot directly control the manager's hidden action, she understands the implications of a contract on the manager's incentives to exert effort. If her conjecture about the manager's hidden action coincides with the action that the manager has incentives to take, the contract is called incentive compatible.

Definition 1.2 (INCENTIVE COMPATIBLE CONTRACT). *A contract*
$\mathcal{C} = ((C_t)_{t \geq 0}, (K_t)_{t \geq 0}, (\eta_t)_{t \geq 0}, (m_t)_{t \geq 0})$ *is incentive compatible if*

$$(m_t)_{t \geq 0} \in \arg \max_{(\hat{m}_t)_{t \geq 0}} \mathrm{E}^{\mathcal{C},(\hat{m}_t)_{t \geq 0}} \left[\int_0^\infty e^{-\delta t} u(C_t + \hat{m}_t \lambda K_t) \, dt \Big| \mathcal{F}_0 \right].$$

The notation $\mathrm{E}^{\mathcal{C},(\hat{m}_t)_{t \geq 0}}[\cdot|\mathcal{F}_0]$ explicitly expresses the fact that the distribution of returns depends on the contract $\mathcal{C}$ and the hidden action strategy $(\hat{m}_t)_{t \geq 0}$. Without loss of generality, we can consider only contracts that are incentive compatible. Because equilibrium strategies are common knowledge, the advisor can always change any given contract to another one in which her conjecture about the manager's hidden action is consistent with the manager's best response to the contract.

Given an incentive-compatible contract, players will symmetrically learn about the skill of the manager by observing the fund's returns. Suppose all players possess a common prior, $\mathrm{E}[h|\mathcal{F}_0] = p \in [0, 1]$. Because the advisor and the investors correctly anticipate the hidden action of the manager, at any time t, all payers have common beliefs about the manager's skill given by

$$\phi_t = \mathrm{E}[h|\mathcal{F}_t].$$

Beliefs are an important state variable of the model. Beliefs determine investors'

expectations about the fund's returns,

$$\mathrm{E}\left[\mu(\eta_t, h, m_t)\,|\,\mathcal{F}_t\right] = \mu(\eta_t, \phi_t, m_t),$$

and, through expected returns, beliefs determine the investor's willingness to supply capital.

Proposition 1.1. *If a contract* $\mathcal{C} = ((C_t)_{t\geq 0}, (K_t)_{t\geq 0}, (\eta_t)_{t\geq 0}, (m_t)_{t\geq 0})$ *is incentive-compatible, then beliefs* $\phi_t = \mathrm{E}[h|\mathcal{F}_t]$ *evolve as*

$$d\phi_t = \eta_t \phi_t (1 - \phi_t) dW_t^{\mathcal{C}}, \tag{1.4}$$

where

$$dW_t^{\mathcal{C}} = \frac{1}{\sigma}\left[dR_t - (r + \mu(\eta_t, \phi_t, m_t))dt\right] \tag{1.5}$$

is an increment to a standard Brownian motion under the measure of returns induced by $\mathcal{C}$.

Equation (1.4) shows the role of experimentation in the production of information. If the advisor requires higher experimentation η_t, the signal-to-noise ratio increases, together with the response of beliefs to any given return shock $dW_t^{\mathcal{C}}$.

Given their beliefs ϕ_t and the contract $\mathcal{C}$, competitive and risk-neutral investors provide capital to the fund through a series of spot-market contracts. These contracts specify the proportional fees that investors pay to the advisor. Because investors are competitive, they take the fund's net expected returns as given and they compare them to the interest rate r, which represents their outside option. Given their risk

17

neutrality, they are not concerned about the idiosyncratic risk of the fund.

Definition 1.3 (SPOT-MARKET CONTRACT). *Given a contract $\mathcal{C}$ and beliefs ϕ_t, a spot-market contract, $\mathcal{C}^S$, is a pair of $\mathcal{F}_t$-measurable variables, (f_t, K_t) such that for all times t investors are willing to supply capital K_t and pay fees f_t,*

$$K_t \in \arg\max_K \ (\mu(\eta_t, \phi_t, m_t) - f_t)K,$$

and such that the advisor obtains revenues $f_t K_t$.

From this definition, we can see that competitive investors offer a perfectly elastic supply of capital at rate r. The advisor can increase fees up to the point at which they coincide with expected excess returns. Given these fees, individual investors are willing to supply any amount of capital, whereas the advisor maximizes the revenues she collects per unit of assets under management. I call a spot market contract optimal if it maximizes the spot revenues of the advisor.

Lemma 1.1. *In any optimal spot-market contract, $f_t = \mu(\eta_t, \phi_t, m_t)$ for all $t \geq 0$.*

Since the proof is standard, I omit it. The lemma is based on the intuition that, if investors are willing to provide at least K_t units of capital and pay a fee $f'_t < \mu(\eta_t, \phi_t, m_t)$, they are also willing to provide at least K_t units of capital for any fee $f_t \in [f'_t, \mu(\eta_t, \phi_t, m_t)]$. A revenue-maximizing advisor therefore chooses the maximum fee investors that are willing to pay. This fee coincides with the expected return $\mu(\eta_t, \phi_t, m_t)$.

Whereas investors' preferences pin down fees f_t, the size of the fund is pinned

18

down by the advisor's demand.[6] By setting fees equal to expected excess returns, the advisor is capturing the value added of the fund, $K_t\mu(\eta_t, \phi_t, m_t)$, through her revenues. She therefore has incentives to design a compensation contract that maximizes the value of the fund.

The compensation contract of the manager will therefore be optimal for the advisor if it minimizes the costs of running the fund, after taking into account the incentives that the manager receives from the contract. Formally, an optimal contract can be defined as follows:

Definition 1.4 (OPTIMAL CONTRACT). $\mathcal{C} = ((C_t)_{t\geq 0}, (K_t)_{t\geq 0}, (\eta_t)_{t\geq 0}, (m_t)_{t\geq 0})$ *is an optimal contract for initial beliefs ϕ_0 and for initial promised value V_0 if*

$$\mathcal{C} \in \arg\inf_{\hat{\mathcal{C}}} \mathrm{E}^{\hat{\mathcal{C}}, (\tilde{m}_t)_{t\geq 0}} \left[\int_0^\infty e^{-rt} \left(\frac{\hat{C}_t}{q(\hat{\eta}_t)} dt - \hat{K}_t\mu(\hat{\eta}_t, h, \tilde{m}_t) \right) dt \middle| \mathcal{F}_0 \right]$$

$$s.t. \quad (\tilde{m}_t)_{t\geq 0} \in \arg\max_{(m'_t)_{t\geq 0}} \mathrm{E}^{\hat{\mathcal{C}}, (m'_t)_{t\geq 0}} \left[\int_0^\infty e^{-\delta t} u(\hat{C}_t + m'_t\lambda\hat{K}_t) dt \middle| \mathcal{F}_0 \right]$$

$$\mathrm{E}^{\hat{\mathcal{C}}, (\tilde{m}_t)_{t\geq 0}} \left[\int_0^\infty e^{-\delta t} u(\hat{C}_t + \tilde{m}_t\lambda\hat{K}_t) dt \middle| \mathcal{F}_0 \right] \geq V_0$$

$$\mathrm{E}[h|\mathcal{F}_0] = \phi_0.$$

In the optimal contract, the advisor explicitly takes into account that the manager might face incentives to shirk and gain private benefits. By shirking, the manager

---

6. The logic would be slightly different in the alternative scenario in which fees are fixed. In this case, the advisor would determine the amount of actively managed assets. Investors would then supply additional capital to be invested in the passive, monitored portfolio that does not involve agency frictions. In this framework, investors determine the total size of the fund because they keep supplying capital until the fund's total net return coincides with the interest rate. As already mentioned, these two scenarios imply identical outcomes for the size of the active portfolio and for the value of the advisor's revenues.

affects expected returns $\mu(\hat{\eta}_t, h, \tilde{m}_t)$, as well as the distribution of returns. Because a contract's terms are written as functions of the history of returns, the expectations in the definition involves the probability measure of returns induced by the contract $\mathcal{C}$ and by the agent's shirking process $(\tilde{m}_t)_{t \geq 0}$. Finally, the optimality of a contract is always defined with respect to a promised value for the manager V_0, which can be seen as an initial outside option for the manager.

In this paper, I am interested in optimal contracts. Although verifying optimality in the sense of Definition 1.4 appears intractable, the following proposition shows that it suffices to search over a restricted class of contracts.

Proposition 1.2. *The optimal contract is incentive compatible with full effort, that is, $m_t = 0$ for all t.*

The intuition for this proposition is the following. First, as already discussed, we lose no generality if we restrict attention to incentive-compatible contracts. Second, optimal contracts must align the manager's incentives to the advisor's objectives. Because shirking is inefficient and the advisor obtains revenues from the fund's value added, it is intuitive that the optimal contract will be designed to implement full effort.

From an operational point of view, Proposition 1.2 asserts that, in searching for an optimal contract, we can simply derive the conditions under which the manager has no incentive to shirk and, within the class of contracts that satisfy this condition, we can select the optimal one. In the remainder of this section, I derive the conditions that an optimal contract must satisfy. I use these conditions in section 1.4 to derive the optimal contract.

1.3.4 Incentive Compatibility and Information Rent

In the previous subsection, I argued that we can restrict our attention to contracts that are incentive compatible and that induce the manager to exert full effort. In this subsection, I provide conditions for a contract to achieve this target. As in static principal-agent problems, these conditions require the manager to be exposed to some fund-level risk. Whereas in static models the principal exposes the agent to risk by giving him a performance-contingent pay, in a dynamic model the principal exposes the agent to risk by adjusting his future continuation value.

To design an optimal contract, the advisor needs to account for the incentives of the manager. At any time t, the manager has incentives to shirk because he obtains consumption value from shirking. However, he could be deterred from shirking if shirking causes a loss in future utility. The manager's future utility coincides with his continuation value V_t, which is a function of the continuation contract[7] and belies, and which can be expressed as

$$V_t = V(\mathcal{C}_t, \phi_t) = \mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)} u(C_s)\, ds \Big| \mathcal{F}_t\right], \qquad (1.6)$$

for an incentive-compatible contract enforcing full effort. The continuation value V_t represents the present value of the future utility that the manager expects to receive from the contract, given his current beliefs ϕ_t.[8]

Using the martingale representation approach developed in previous literature

---

7. A continuation contract a time t, $\mathcal{C}_t$, is a set of $\mathcal{F}_s^t$-adapted processes $((C_s)_{s\geq t}, (K_s)_{s\geq t}, (\eta_s)_{s\geq t}, (m)_{s\geq t})$, where $\mathcal{F}_s^t = \{(R_u)_{t\leq u\leq s}\}$.

8. Given a continuation contract $\mathcal{C}_t$, beliefs are a sufficient statistic for the probability measure

(Sannikov, 2008; Williams, 2009), I obtain the law of motion for the manager's continuation value V_t.

Lemma 1.2. *The manager's continuation value evolves as*

$$dV_t = (\delta V_t - u(C_t))dt + \beta_t dW_t^{\mathcal{C}} \tag{1.7}$$

for some $\mathcal{F}_t$-adapted process $(\beta_t)_{t \geq 0}$.

The proof of Lemma 1.7 is standard and it is therefore omitted.

If β_t is different from zero, the manager is facing a risky consumption path, which is inefficient. In a fully efficient allocation, the risk-neutral investors and advisor should fully insure the risk-averse manager. However, because of moral hazard, the advisor designs a performance-based contract $\mathcal{C}$ that exposes the manager to some risk in order to provide incentives. If β_t is positive, the utility of the manager increases if he delivers a return that exceeds expectations. If the manager shirks, expected returns decline and the manager suffers a loss of future utility.

As a benchmark, assume for the moment that the skill of the manager, h, is known. In this case, the advisor can prevent shirking by offering a contract in which the manager's exposure to returns, β_t, offsets the marginal consumption value of shirking.

---

of future returns. To see why, write

$$V_t = (1 - \phi_t)\mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}u(C_s)\,ds\middle|h = 0\right] + \phi_t\mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}u(C_s)\,ds\middle|h = 1\right].$$

The conditional expectations on the right-hand side of this equation are functions of the continuation contract $\mathcal{C}_t$ only.

22

Lemma 1.3. *If h is common knowledge, a necessary and sufficient condition for incentive compatibility is*

$$u'(C_t)\lambda\sigma K_t \leq \beta_t.$$

A proof for Lemma 1.3 can be found in Di Tella (2017) and Di Tella and Sannikov (2018).

The condition in Lemma 1.3 is intuitive. If the manager shirks, he reduces returns by m_t, and hence $dW_t^{\mathcal{C}}$ acquires a negative drift $-\frac{m_t}{\sigma}$. The manager therefore suffers a loss of continuation value equal to $\beta_t \frac{m_t}{\sigma}$. However, his current utility increases from $u(C_t)$ to $u(C_t + m_t\lambda K_t)$. The condition in Lemma 1.3 ensures that $m_t = 0$ is the best response of the manager to the contract, that is,

$$0 = \arg\max_{m\geq 0} \quad \underbrace{u(C_t + m\lambda K_t)}_{\text{benefit of shirking}} - \underbrace{\beta_t \frac{m}{\sigma}}_{\text{cost of shirking}}.$$

If the skill of the manager is uncertain, the advisor faces some additional challenges in designing an incentive-compatible contract. For example, suppose the manager deviates to $m_t = m > 0$ for a small amount of time between s and $s+\Delta s$. With learning, the manager not only gains consumption value from the deviation, but he also earns an informational advantage over the advisor and the investors. Unaware of the manager's deviation, the advisor and the investors update beliefs according to equations (1.4) and (1.5). Because the true drift of returns is now $\mu(\eta_t, h, m)$, their beliefs, ϕ_t, will acquire a negative drift relative to the manager's beliefs, $\tilde{\phi}_t$. Immediately after the deviation, the difference between the manager's and the other

players' beliefs, $\tilde{\phi}_t - \phi_t$, will be given by

$$\tilde{\phi}_{s+\Delta s} - \phi_{s+\Delta s} \approx \eta_s \phi_s (1 - \phi_s) \frac{m}{\sigma} \Delta s.$$

This difference in beliefs is persistent and causes persistent distortions in the provision of incentives. Following a deviation, the manager will always be more optimistic than the other players, that is, $\tilde{\phi}_t > \phi_t$ for all $t > s + \Delta s$. The manager is not only more optimistic, but also aware of possessing correct beliefs. By having more accurate and optimistic beliefs, the manager earns an information rent over the other players. Intuitively, a skilled manager who is believed to be unskilled is better off than a manager who is actually unskilled. The skilled managers can expect to surprise the market in the future thanks to his superior skill. The truly unskilled manager cannot expect to surprise anyone. To see how a manager earns an information rent, consider equation (1.7). If $\tilde{\phi}_t > \phi_t$, shocks $\frac{1}{\sigma} [dR_t - (r + \mu(\eta_t, \phi_t, 0))dt]$ have a positive drift given by

$$\eta_t(\tilde{\phi}_t - \phi_t)dt > 0.$$

Therefore, the manager's continuation value acquires a positive drift

$$\beta_t \eta_t (\tilde{\phi}_t - \phi_t)dt.$$

The drift $\beta_t \eta_t (\tilde{\phi}_t - \phi_t)$ captures the surprise that the manager expects other players to receive. Suppose that, after a (hidden) deviation, the advisor promises a continuation value $V_{s+\Delta s}$ to the manager. However, the true continuation value of the manager

is larger than what the advisor explicitly promises, and it includes the additional surplus that accrues to the manager through the future drift $\beta_t \eta_t (\tilde{\phi}_t - \phi_t)$. The present value of this additional surplus constitutes the information rent that the manager has over the advisor and investors.

This reasoning suggests that when players are learning, an incentive-compatible contract should provide incentives β_t to offset the manager's marginal utility of shirking as well as the information rent that he could earn. I formalize this argument using the stochastic maximum principle introduced in the contract-design literature by Williams (2011).

Proposition 1.3. *In any optimal contract,*

$$u'(C_t)K_t\lambda\sigma \leq \beta_t - \eta_t\xi_t, \tag{1.8}$$

where ξ_t follows

$$d\xi_t = (\delta\xi_t - \eta_t\beta_t\phi_t(1 - \phi_t))dt + \omega_t dW_t^{\mathcal{C}} \tag{1.9}$$

for some $\mathcal{F}_t$-adapted process $(\omega_t)_{t\geq 0}$.

The term $\eta_t\xi_t$ in the incentive-compatibility condition (1.8) accounts for the information rent that accrues to the manager after a deviation. To interpret the variable ξ_t more clearly, consider expression (1.6). In particular, consider the fact that at any point in time, the manager's continuation value is a function of the continuation contract $\mathcal{C}_t$ and his beliefs ϕ_t, that is $V_t = V(\mathcal{C}_t, \phi_t)$. For a given contract, $\partial_\phi V(\mathcal{C}_t, \phi_t)$ measures the marginal change in continuation value coming from a marginal change in beliefs. The variable ξ_t is related to this marginal of beliefs, and therefore can be

25

defined as the information rent of the manager.

Proposition 1.4.

$$\xi_t = \phi_t(1 - \phi_t)\partial_\phi V(\mathcal{C}_t, \phi_t). \tag{1.10}$$

Moreover,

$$\xi_t = \mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}\eta_s\beta_s\phi_s(1 - \phi_s)\,ds\bigg|\mathcal{F}_t\right], \tag{1.11}$$

and in any incentive-compatible contract, $\xi_t \geq 0$.

I can now give an intuitive interpretation of the incentive-compatibility condition (1.8). First, from equation (1.7), we know that incentives β_t correspond to the total volatility of the continuation value V_t. Second, combining (1.4) and (1.10), we see that $\eta_t\xi_t = \eta_t\phi_t(1-\phi_t)\partial_\phi V(\mathcal{C}_t, \phi_t)$ represents the volatility of the continuation value that originates from the changes in beliefs. Therefore, we can interpret the quantity

$$\beta_t - \eta_t\xi_t = \beta_t - \eta_t\phi_t(1 - \phi_t)\partial_\phi V(\mathcal{C}_t, \phi_t)$$

as the volatility of the manager's continuation value that originates from changes in the continuation contract while keeping beliefs fixed.

Seen from this perspective, the incentive-compatibility condition (1.8) is extremely intuitive. Equation (1.8) states that, to provide incentives to the manager, the advisor cannot rely on changes in beliefs to punish him for bad performance. Although changes in beliefs do affect the volatility of the continuation value along the equilibrium path, they cannot be exploited to prevent off-equilibrium deviations. The reason is that the manager's deviations do not affect his beliefs. After a devi-

26

ation m_t, the manager's true expected future utility declines by $(\beta_t - \eta_t \xi_t) \frac{m_t}{\sigma}$, and not by $\beta_t \frac{m_t}{\sigma}$, as the advisor incorrectly thinks. Therefore, in an incentive-compatible contract, the quantity $\beta_t - \eta_t \xi_t$ is what matters for incentive provisions, and must be such that

$$0 = \arg\max_{m \geq 0} \quad \underbrace{u(C_t + m\lambda K_t)}_{\text{benefit of shirking}} - \underbrace{(\beta_t - \eta_t \xi_t) \frac{m}{\sigma}}_{\text{cost of shirking}}.$$

In other words, the incentive-compatibility condition asserts that, to provide incentives to the manager, the advisor must adjust the continuation contract so that, keeping beliefs constant, the incentives of the manager exceed the private benefit of shirking.

With learning, the advisor faces additional costs in providing incentives to the manager, as indicated by the result in Proposition 1.4 that the information rent ξ is positive. According to equation (1.8), for given consumption C_t, capital K_t, and experimentation η_t, a manager with uncertain skill will have to bear an additional quantity $\eta_t \xi_t$ of risk relative to a manager with known skill. Because the manager is risk averse, the advisor will have to compensate him for this additional risk with a higher expected compensation, thus increasing the cost of the contract.

Any contract implies an information rent for the manager, and a larger information rent implies higher costs for the advisor. Therefore, from an ex ante perspective, the advisor prefers to design a contract that implies a small information rent. To do so, she has to commit to reduce incentives, β_s, and experimentation, η_s, in the future because, as equation (1.11) shows, the manager's information rent ξ_t is a present value of these quantities. However, from an ex post perspective, lowering incentives

and experimentation may not be optimal. Consequently, in order to achieve ex ante optimality, the advisor needs to fully commit to the terms of an initial contract. In sections 1.4 and 1.5, I study the optimal contract with full commitment and discuss how the trade-off between ex ante and ex post optimality affects the flow-performance relationship and managerial compensation.

1.3.5   Verifying Incentive Compatibility

I conclude this section by presenting a condition that can be used to verify the incentive compatibility of a contract. Proposition 1.3 offers a condition that prevents the manager from engaging in a one-shot deviation. Although this condition is necessary for incentive compatibility, alone it does not guarantee full incentive compatibility. With learning, any hidden shirking will cause a persistent wedge between the manager's and the other player's beliefs. Given this wedge, the condition in Proposition 1.3 may not be sufficient to prevent shirking. Even if (1.8) holds, the manager's best response to the contract may still involve a dynamic shirking strategy.

Previous literature has long recognized the challenge that state variables, like beliefs, pose in the design of an optimal contract. The common approach is to solve for an optimal contract by imposing the necessary condition (1.8) only. This approach is the so called relaxed-problem approach. Then, one should verify whether the contract so obtained satisfies a sufficient incentive-compatibility condition. This strategy is the one undertaken by He (2012) and Di Tella and Sannikov (2018) for private savings, by Prat and Jovanovic (2014), DeMarzo and Sannikov (2017), He et al. (2017), and Cisternas (2018) for learning, and Williams (2011) for generic

state variables. I take the same approach and use the following proposition to verify whether the solution to the relaxed problem is actually incentive compatible.

Proposition 1.5. *If*

$$u'(C_t)K_t\lambda\sigma + \eta_t\xi \leq \beta_t$$

and

$$\omega_t \geq \eta_t(1 - 2\phi_t)\xi_t, \tag{1.12}$$

then the contract is incentive compatible.

This proposition contains a sufficient condition for incentive-compatibility. In general, incentive-compatible optimal contracts must satisfy (1.8), but not necessarily (1.12). In solving for an optimal contract, I adopt the first-order approach and impose the necessary condition (1.8) as a constraint on the contract. It can be later verified whether such a contract satisfies condition (1.12).

To interpret the sufficient condition for incentive compatibility, it is useful to refer to the proof of this proposition in Appendix 1.C, where I show that $\omega_t - \eta_t(1-2\phi_t)\xi_t$ is proportional to the volatility of $\partial_\phi V(\mathcal{C}_t, \phi_t)$. Proposition 1.5 therefore states that if a contract prevents instantaneous deviations and it reduces the marginal value of beliefs after a negative shock, then it is a fully incentive-compatible contract.

This result has some intuitive appeal. If the contract lowers the marginal value of the manager's beliefs $\partial_\phi V(\mathcal{C}_t, \phi_t)$ after a bad shock then, following a deviation, the manager would suffer a decrease not only in his continuation value, but also in his information rent. The manager loses part of the option to "impress" other players in the future, thus lowering the value of his informational advantage. Together, this

29

condition and condition (1.8) are sufficient to induce the manager to always exert full effort.

Equation (1.12) is likely to hold in an optimal contract. Looking at Proposition 1.4, we see ξ_t depends on future incentives β_s and experimentation η_t. After a good shock, the advisor has incentives to increase future incentives and future experimentation because of the following reasons. First, a good shock increases expected returns and the advisor will likely take advantage of them by increasing capital under management K_t and, by (1.8), incentives β_t. Second, experimentation becomes more profitable, so the advisor will likely increase experimentation η_t as well. We therefore have reasonable economic motivations to believe that in the optimal contract, ξ_t increases after good performance.

1.4   Optimal Contract

Given the necessary incentive-compatibility condition in Proposition 1.3, I adopt a first-order approach to solve for the optimal contract under full commitment. According to the first-order approach, I solve for the optimal contract in Definition 1.4 as a recursive problem subject to the incentive-compatibility condition (1.8) at every point in time. The advisor is fully committed to the manager's continuation value V_t and to his information rent ξ_t, which therefore constitute the recursive state variables of the problem together with beliefs ϕ_t[9]. With full commitment, the advisor always

---

9. Although full commitment from both the manager and the advisor is certainly a strong assumption, this formulation of the problem offers a key benchmark for alternative specifications. With full commitment, the advisor implements an allocation that yields the best outcome given the frictions of the model. After relaxing the full-commitment assumption in Appendix 1.A, I highlight which results hold independently of the contracting assumptions and which results depend instead

honors her past promises in terms of continuation value and information rent, and the laws of motion (1.7) and (1.9) represent promise-keeping constraints for the advisor. To provide incentives, the advisor specifies how future continuation values and information rents evolve on the basis of performance. She therefore selects incentives β_t and volatility ω_t optimally.

Formally, the optimal contract is characterized as a solution to the following optimization problem:

$$J^*(V_0, \xi_0, \phi_0) = \inf_{(C_t, K_t, \beta_t, \omega_t, \eta_t)_{t \geq 0}} \mathrm{E}\left[\int_0^\infty e^{-rt} \left(\frac{C_t}{q(\eta_t)} - K_t \mu(\eta_t, \phi_t, 0) \right) dt \,\bigg|\, \mathcal{F}_0 \right]$$

$$\text{s.t.} \quad C_t^{-\rho} \lambda \sigma K_t + \eta_t \xi_t \leq \beta_t \quad \forall t \geq 0$$

$$dV_t = \left(\delta V_t - \frac{C^{1-\rho}}{1-\rho} \right) dt + \beta_t dW_t^{\mathcal{C}}$$

$$d\xi_t = \left(\delta \xi_t - \eta_t \beta_t \phi_t (1 - \phi_t) \right) dt + \omega_t dW_t^{\mathcal{C}}$$

$$d\phi_t = \eta_t \phi_t (1 - \phi_t) dW_t^{\mathcal{C}}.$$

(1.13)

Problem (1.13) is clearly challenging to solve analytically. However, we can characterize some of its properties. For this purpose, consider its associated Hamilton-

---

on the particular contractual form that is implemented.

Jacobi-Bellman (HJB) equation,

$$
\begin{aligned}
rJ^*(V,\xi,\phi) = \inf_{\substack{C,K,\beta,\omega,\eta \\ C^{-\rho}\lambda\sigma K+\eta\xi\leq\beta}} & \left\{ \frac{C}{q(\eta)} - (\alpha+\sigma\eta\phi)K \right. \\
& + J_V^*(V,\xi,\phi)\left(\delta V - \frac{C^{1-\rho}}{1-\rho}\right) + J_\xi^*(V,\xi,\phi)\left(\delta\xi - \eta\beta\phi(1-\phi)\right) \\
& + \frac{1}{2}J_{VV}^*(V,\xi,\phi)\beta^2 + \frac{1}{2}J_{\xi\xi}^*(V,\xi,\phi)\omega^2 + \frac{1}{2}J_{\phi\phi}^*(V,\xi,\phi)\eta^2\phi^2(1-\phi)^2 \\
& \left. + J_{V\xi}^*(V,\xi,\phi)\beta\omega + J_{V\phi}^*(V,\xi,\phi)\beta\eta\phi(1-\phi) + J_{\xi\phi}^*(V,\xi,\phi)\omega\eta\phi(1-\phi) \right\}. \quad (1.14)
\end{aligned}
$$

First, we can verify that, in any optimal contract, the incentive-compatibility condition holds with equality, that is,

$$
K_t = \frac{\beta_t - \eta_t\xi_t}{\lambda\sigma}C_t^\rho. \quad (1.15)
$$

Because excess returns and fees are always positive, the advisor increases the level of assets under management as much as the incentive-compatibility constraint permits. Investors do not provide any capital beyond this level at the current fees because if the incentive-compatibility condition is violated, the manager will shirk and expected returns will not cover the fees investors pay.

Second, the HJB equation (1.14) is subject to a boundary condition $J^*(V,0,\phi) = J^0(V)$, where the function $J^0(V)$ is the cost function for the advisor when the man-

ager lacks skill, that is, when $h = 0$ is common knowledge.[10] This boundary condition is motivated by the promise-keeping constraint on the information rent ξ. If the advisor promises zero information rent to the manager, the only way she can keep this promise is by providing no incentives β, and hence no capital K, and/or by stopping experimentation η forever. The advisor clearly prefers to stop experimentation only, because she can still obtain fees equal to the baseline excess return α by providing capital to the manager.

We cannot obtain analogous boundary conditions for $J^*(V, \xi, 0)$ and $J^*(V, \xi, 1)$. The contract never reaches these boundaries in finite time, because beliefs ϕ have no drift and their volatility vanishes as they approach 0 and 1. These singular points, however, do not constitute a problem. We can indeed think of the HJB equation (1.14) as a state constraint problem whereby beliefs are constrained between 0 and 1. Katsoulakis (1994) and Alvarez et al. (1997) show that state constraints effectively replace boundary conditions in determining the solution of partial differential equations.

Finally, we can derive an endogenous bound for the information rent ξ. Note that at time 0, the advisor has no initial commitment to any information rent. As illustrated in Definition 1.4, the optimal contract is initially defined only in terms of the initial promised value of the manager, V_0, and initial beliefs, ϕ_0. Therefore, we can think of the advisor as first setting an initial information rent ξ_0, and then solving

---

10. $J^0(V)$ satisfies the HJB equation

$$rJ^0(V) = \inf_{C,\beta} \left\{ C - \alpha\frac{\beta}{\sigma\lambda}C^\rho + J_V^0(V)\left(\delta V - \frac{C^{1-\rho}}{1-\rho}\right) + \frac{1}{2}J_{VV}^0(V)\beta^2 \right\}.$$

problem (1.13) given the chosen ξ_0. The advisor chooses an initial information rent that minimizes her costs. If $J_\xi^*(V, \xi, \phi)$ is convex in ξ and a global minimum exists, the advisor will pick ξ_0 such that

$$J_\xi^*(V_0, \xi_0, \phi_0) = 0.$$

We can further characterize the behavior of the information rent and its marginal cost J_ξ^* in the optimal contract. For any pair of continuation value V and beliefs ϕ, define $\bar{\xi}(V, \phi)$ as the information rent that minimizes costs for the advisor. I formally prove the following proposition in Appendix 1.D. However, one could also make some convexity and differentiability assumptions on $J^*(V, \xi, \phi)$ and directly use the envelope theorem on (1.14) as in DeMarzo and Sannikov (2017) to derive the following result.

Proposition 1.6. *In the optimal contract, two properties hold:*

I. The marginal cost of information rent is always non-positive,

$$J_\xi(V_t, \xi_t, \phi_t) \leq 0 \quad t \geq 0$$

II. For all $t \geq 0$,
$$\xi_t \leq \bar{\xi}(V_t, \phi_t).$$

Information rents in contracting models tend to be bounded, and propositions analogous to 1.6 have been derived in DeMarzo and Sannikov (2017) and He et al.

(2017). Large information rents expose the manager to risks that are irrelevant for incentive provision. Consequently, the advisor avoids promising excessive information rents.

Even after deriving the properties that I have discussed so far, numerically solving the HJB equation (1.14) is far from easy, if not far from feasible. Proposition 1.6 provides the reason. In the optimal contract, the information rent is bounded by $\bar{\xi}(V, \phi)$. For values of information rent far above this bound, optimal contracts might not exist, or they might violate regularity properties required for the validity of a dynamic programming approach. To solve for the optimal contract, we need to know the bound $\bar{\xi}(V, \phi)$. However, to derive this bound, we need to know the optimal contract. We could attempt to numerically solve (1.14) by sequentially guessing and verifying the bound function $\bar{\xi}(V, \phi)$. This attempt would be extremely computationally inefficient, if not unfeasible. I take a different approach.

Proposition 1.6 itself offers a hint to solve the model efficiently. In the optimal contract, the marginal cost $J_\xi(V_t, \xi_t, \phi_t)$ is bounded above by zero, thus always satisfying the endogenous bound on the information rent. I then formulate a problem in which the marginal cost $J_\xi(V_t, \xi_t, \phi_t)$ replaces the information rent ξ as a state variable. This problem is connected to the initial one through a duality relation. In the remainder of this section, I introduce the dual problem of (1.13), derive its properties, and show that the dual problem offers an efficient way to characterize the contract.

1.4.1   The Dual Problem

I now introduce the dual problem of (1.13). Informally speaking, the purpose of the dual problem is to replace the information rent, ξ_t, with its marginal value, $J_\xi(V_t, \xi_t, \phi_t)$, as a state variable.

Define the multiplier

$$
Y_t = e^{t(r-\delta)}\left[-\left(\int_0^t e^{s(\delta-r)}\mu(\eta_t, \phi_t, 0)\frac{\eta_t}{\lambda\sigma}C_t^\rho\, ds\right) + Y_0\right],
$$

where $Y_0 \in \mathbb{R}$, and consider the following problem:

$$
G^*(V_0, Y_0, \phi_0) =
$$

$$
\inf_{(C_t,\beta_t,\eta_t),\forall t\geq 0} \mathrm{E}\left[\int_0^\infty e^{-rt}\left(\frac{C_t}{q(\eta_t)} - \mu(\eta_t, \phi_t, 0)\frac{\beta_t}{\lambda\sigma}C_t^\rho - Y_t\eta_t\beta_t\phi_t(1-\phi_t)\right) dt\,\Big|\mathcal{F}_0\right]
$$

$$
\text{s.t.}\quad dV_t = (\delta V_t - u(C_t))dt + \beta_t dW_t^\mathcal{C}
$$

$$
dY_t = (r-\delta)Y_t dt - \mu(\eta_t, \phi_t, 0)\frac{\eta_t}{\lambda\sigma}C_t^\rho dt
$$

$$
d\phi_t = \eta_t\phi_t(1-\phi_t)dW_t^\mathcal{C}.
$$

$$
(1.16)
$$

I call (1.16) the dual problem, as opposed to problem (1.13), which, hereafter, I call the primal problem.

The dual problem in (1.16) has an intuitive appeal. Consider the objective function. The term $\frac{C_t}{q(\eta_t)} - \mu(\eta_t, \phi_t, 0)\frac{\beta_t}{\lambda\sigma}C_t^\rho$ represents the flow cost of an advisor who runs a fund with no information-rent problem. We can think of this hypothetical fund as one in which the hidden action m_t is observable but not contractible. In

this case, the incentive-compatibility condition is the same as in Lemma 1.3, because shirking does not induce any belief distortion or any information rent for the manager. The term $-Y_t\eta_t\beta_t\phi_t(1-\phi_t)$ represents a penalty for an advisor who sets large incentives β_t or large experimentation rate η_t. The severity of the penalty is measured by the multiplier Y_t. The advisor faces a larger penalty for incentives and experimentation if the multiplier Y_t is more negative. In the dual problem (1.16), the advisor replaces the information rent of the manager with the multiplier Y as a relevant state variable.

Having introduced the dual problem, I can finally show the connection with the primal problem (1.13) and illustrate how the optimal contract can be derived as a solution to the dual problem (1.16).

Proposition 1.7. *The optimal contract is the solution to the dual problem* (1.16) *with $Y_0 = 0$. The primal cost function J^* and the dual cost function G^* are related by*

$$J^*(V_t, \xi_t, \phi_t) = \sup_{Y \leq 0} \{G^*(V_t, Y, \phi_t) + Y\xi_t\}. \tag{1.17}$$

At any time t, the optimal contract implies an information rent $\xi_t = -G_Y^*(V_t, Y_t, \phi_t)$. Moreover, $Y_t = J_\xi^*(V_t, \xi_t, \phi_t)$, and $Y_t \leq 0$ for all $t \geq 0$.

Proposition 1.7 is extremely powerful. It states that any optimal contract can be obtained as a solution to the dual problem if the initial multiplier is chosen appropriately. Moreover, by combining Proposition 1.7 with Proposition 1.6, we obtain a recursive characterization of the optimal contract through a dual formulation that automatically satisfies the endogenous bounds on the information rent.

37

This proposition is key to overcome the computational challenges that I discussed before introducing the dual problem. Looking at the law of motion of the multiplier Y_t in (1.16), we see that, for any choice of $(C_t, \beta_t, \eta_t)_{t \geq 0}$, Y_t always be non-positive as long as $Y_0 \leq 0$. Because $Y_t = J_\xi^*(V_t, \xi_t, \phi_t)$, if I restrict the state space to non-positive value of the multiplier Y, then the marginal cost of information rent is guaranteed to remain non-positive. Therefore, standard numerical methods are sufficient to obtain a solution for the optimal contract.

Before numerically solving the dual problem, I exploit the homogeneity of the problem and introduce some notation that will simplify the numerical computations and the discussion of the results. On the basis of Lemma 1.10 in Appendix 1.D, I can simplify the numerical burden of solving for problem (1.16) by reducing the number of state variables from three to two. In particular, we can write the dual cost function as

$$G^*(V, Y, \phi) = \hat{v} g^*(y, \phi),$$

where

$$\hat{v} = ((1 - \rho)V)^{\frac{1}{1-\rho}}$$

is the consumption equivalent of the manager's continuation value, and where

$$y = (1 - \phi)\hat{v}^{-\rho}Y$$

is a scaled version of multiplier Y. We can then define the scaled control variables

$$c_t = \frac{C_t}{\hat{v}_t}, \quad k_t = \frac{K_t}{\hat{v}_t}, \quad \text{and} \quad \hat{\beta}_t = \frac{\beta_t}{(1 - \rho)V_t},$$

and derive the law of motion of continuation value $\hat{v}$ and multiplier y through Ito's lemma, thus obtaining

$$\frac{d\hat{v}_t}{\hat{v}_t} = \left(\frac{\delta}{1-\rho} - \frac{c_t^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_t^2 \right) dt + \hat{\beta}_t \hat{v}_t dW_t^{\mathcal{C}} \tag{1.18}$$

and

$$dy_t = -(1-\phi_t)\mu(\eta_t, \phi_t, 0)\frac{\eta_t}{\lambda\sigma}c_t^{\rho} dt$$

$$+ y_t \left(r - \frac{\delta}{1-\rho} + \rho\frac{c_t^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_t^2 + \rho\eta_t\hat{\beta}_t\phi_t \right) dt - y_t[\rho\hat{\beta}_t + \eta_t\phi_t]dW_t^{\mathcal{C}}. \tag{1.19}$$

In conclusion, instead of solving for $G^*(V, Y, \phi)$ as a function of three state variables, I solve the following HJB equation, which characterizes $g^*(y, \phi)$ as a function of the multiplier y and belief ϕ:

$$rg^*(y, \phi) = \inf_{c,\hat{\beta},\eta} \left\{ \frac{c}{q(\eta)} - (\alpha + \sigma\eta\phi)\hat{\beta}\frac{c^{\rho}}{\sigma\lambda} - y\hat{\beta}\eta\phi \right.$$

$$+ g^*(y, \phi) \left(\frac{\delta}{1-\rho} - \frac{c^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}^2 \right)$$

$$+ g_y^*(y, \phi) \left[-(1-\phi_t)\mu(\eta, \phi, 0)\frac{\eta}{\lambda\sigma}c^{\rho} \right.$$

$$\left. +y \left(r - \frac{\delta}{1-\rho} + \rho\frac{c^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}^2 + \rho\eta\hat{\beta}\phi \right) \right] \tag{1.20}$$

$$- g_y^*(y, \phi)y[\rho\hat{\beta} + \eta\phi]\hat{\beta} + g_\phi^*(y, \phi)\eta\phi(1-\phi)\hat{\beta}$$

$$+ \frac{1}{2}g_{yy}^*(y, \phi)y^2[\rho\hat{\beta} + \eta\phi]^2 + \frac{1}{2}g_{\phi\phi}^*(y, \phi)\eta^2\phi^2(1-\phi)^2$$

$$\left. - g_{y\phi}^*(y, \phi)y[\rho\hat{\beta} + \eta\phi]\eta\phi(1-\phi) \right\}.$$

In the next section, I numerically solve the HJB equation (1.20) and characterize the implications of the optimal contract for the fund flows and the manager's compensation. In interpreting the results, readers may refer to the HJB equation (1.20), to the dynamics of promised value (1.18), and to the dynamics of the multiplier y (1.19). However, in my interpretation of the results, I mostly rely on economic intuition and make only minimal references to the specific details of these equations.

1.5 Results and Discussion

I numerically solve the partial differential equation (1.20) by using a finite difference method. To solve for the optimal contract, I use Proposition 1.7 and restrict the state space to negative values of the multiplier y. Because at $y = 0$ the drift of multiplier y is negative and its volatility is 0, I do not need to impose any restrictions on the control variables to satisfy the state constraint $y \leq 0$.

1.5.1 Calibration

To select the parameters of the model, I rely on three main strategies. If a parameter can be directly observed, I use empirical observations of that parameter value. If a parameter cannot be directly observed, but has a strong connection to an outcome of my model that can be observed in the data, I select values of the parameter that yield outcomes that match the data. Finally, I discipline unobservable preference parameters by selecting values previously adopted in the contracting literature. Table 1.1 summarizes the parameter choice.

I set the interest rate r equal to the average real rate starting from the year 2000.

Table 1.1: Model Parameters

Parameter	Notation	Value	Source/Target
Interest rate	r	2.6%	Average real rate from 2000
Baseline excess return	α	0.1%	Minimum fee in Pastor et al. (2015)
Baseline volatility	σ	18%	High vol funds, Pastor et al. (2015)
Discount rate of manager	δ	5%	Di Tella and Sannikov (2018), DeMarzo et al. (2012)
Risk aversion / IES^{-1}	ρ	1/3	Di Tella and Sannikov (2018)
Bound on learning	$\bar{\eta}$	1%	} High alpha funds, Fama and French (2010)
Cost of learning	$\bar{q}$	0.15	
Curvature of learning cost	d	2.1	Convexity of cumulative flows
Gains from shirking	λ	0.85	Flow-performance slope

To select a value for the baseline excess return, α, I consider the sample of mutual funds that I select in section 1.6. As in Pastor et al. (2015), I exclude all funds that charge annual fees below 0.1%, because they are unlikely to be actively managed. I therefore take this value to represent the minimum excess return that a fund is able to provide.

Due to the homogeneity of the model, the volatility of returns must be large enough to guarantee the existence of a finite solution for the contracting problem. Intuitively, if the volatility of returns is too low, the advisor could easily detect shirking. She could therefore easily provide incentives to the manager and could exploit the constant returns to scale of the technology to gain unbounded revenues. By adding decreasing returns in actively managed assets, as in Berk and Green (2004), I could avoid the possibility of unbounded returns and would therefore have additional degrees of freedom in the choice of parameters. Unfortunately, decreasing returns would complicate the problem substantially, because the homogeneity property of the dual cost function G^* would fail. I therefore choose a value of the volatility σ of 18%, which is close to, but a little larger than, the highest percentile of the

standard deviation of abnormal returns in Pastor et al. (2015), which is 16.4% on an annualized basis.

I choose the manager's preference parameters to match choices made in previous contracting literature. I set the discount rate of the manager, δ, equal to 5% as in DeMarzo et al. (2012) and Di Tella and Sannikov (2018), while for the inverse of the manager's IES, I select 1/3 as in Di Tella and Sannikov (2018).

I assume the experimentation cost is a power function of experimentation, that is, $q(e) = 1 - \bar{q}e^d$ for some parameters $\bar{q}$ and d. I set d to 2.1 to obtain sufficient convexity of cumulative flows in cumulative performance. I then consider the distribution of four-factor gross alpha in Fama and French (2010). They find that the 90th percentile of the distribution is 1.3%. I take this number to represent the excess returns of highly-skilled managers. I therefore set $\bar{q} = 0.15$ and $\bar{\eta} = 1\%$ so that a fund with $\phi = 1$ and multiplier $y = 0$ generates expected of 1.3%, and so that the constraint $\eta \leq \bar{\eta}$ does not bind.

Finally, I set the gains from shirking, λ, to 0.85. In this way, the slope of the flow-performance relationship for funds with $\phi = 0.5$ and with multiplier $y \in [-0.1, 0]$ approximately matches the value of 17% estimated in Table 1.4.

1.5.2   Flows and Performance

I define the slope of the flow-performance relationship, ϵ_K, as the percentage change in assets under management for a 1% return, that is,

$$\epsilon_K = \frac{dK_t/K_t}{dR_t}.$$
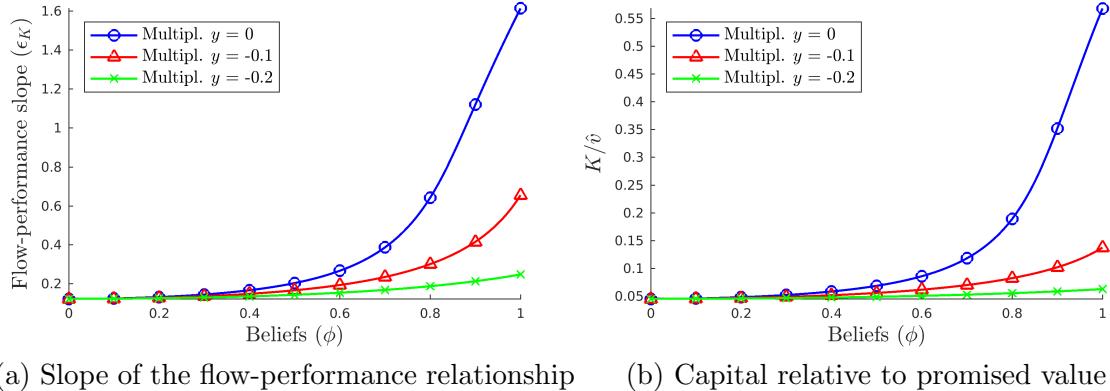
42

(a) Slope of the flow-performance relationship   (b) Capital relative to promised value

Figure 1.1: Slope of the flow-performance relationship and capital in the optimal contract. The parameters of the model are as in Table 1.1.

Using the fact that in the optimal contract, $K_t = \hat{v}_t k(y_t, \phi_t)$ for an optimal capital-to-value ratio $k(y_t, \phi_t)$, the slope of the flow-performance relationship can be measured as

$$\epsilon_K(y, \phi) = \frac{1}{\sigma} \left(\frac{\sigma_k(y, \phi)}{k(y, \phi)} + \hat{\beta}(y, \phi) \right), \tag{1.21}$$

where $\sigma_k(y, \phi)$ is the volatility of $k(y, \phi)$.

As Figure 1.1a shows, the slope of the flow-performance relationship ϵ_K is positive, meaning that assets under management increase after good performance. This result is consistent with several empirical results.

The prediction that sets my model apart from previous theoretical literature, such as Berk and Green (2004), is that the slope of the flow-performance relationship increases after good performance. We can see this result in Figure 1.1a, where the slope of the flow-performance relationship, ϵ_K, increases when the posterior ϕ and the multiplier y increase. These two state variables have positive volatility (recall

43

(a) Fixed fees   (b) Variable fees
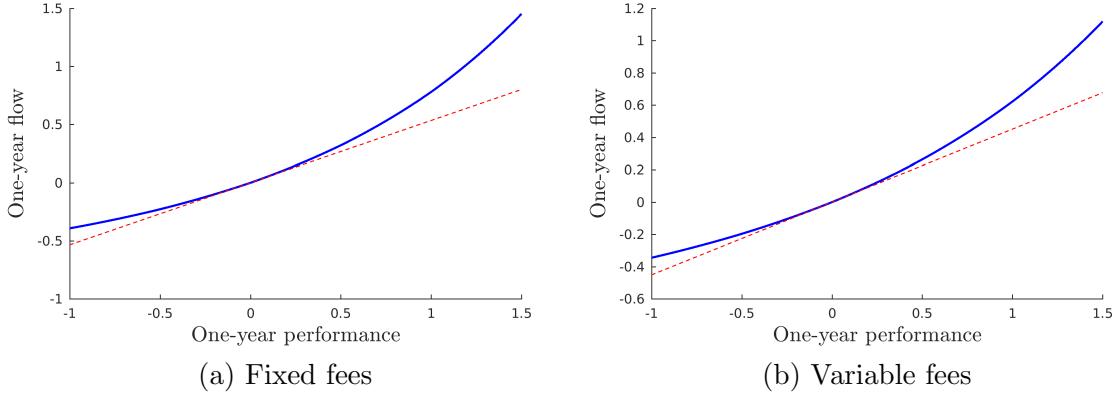
Figure 1.2: Convex relation between cumulative flows and cumulative returns. The solid blue curve represents the cumulative flows as a function of cumulative performance. The dashed red line is the tangent line at 0. Curves are shifted to represent the flows relative to a fund that has a zero cumulative performance. Performance and flows are computed over one year. In 1.2a, fees are assumed to be fixed and flows include changes in actively managed capital, as well as changes in the holdings of a passive, agency-free index. In 1.2b, fees match expected returns, and flows include only changes in actively managed capital. Figures are drawn for initial beliefs $\phi_0 = 0.75$ and initial multiplier $y_0 = 0$. The parameters of the model are as in Table 1.1.

equations (1.4) and (1.19), and recall that $y \leq 0$). Therefore, the slope of the flow-performance relationship ϵ_K increases after good performance. This prediction is new to the literature. Because no other theoretical or empirical paper has studied how the flow-performance relationship depends on the history of a fund's returns, I test this prediction in Section 1.6 using mutual fund data.

Because the slope of the flow-performance relationship ϵ_K increases in past performance, cumulative flows are a convex function of cumulative performance. As an illustration, Figure 1.2 shows the one-year flow into a fund with $y = 0$ and $\phi = 0.75$ as a function of the one-year cumulative return. To make the example more empirically relevant, in Figure 1.2a, I assume the fund charges a fixed fee and collects

additional capital to be invested in a costless, agency-free index. We can clearly observe that the net flows into the fund are a convex function of the cumulative returns. As good performance accumulates over time, additional positive returns have a stronger impact on additional flows. As a result, we observe convexity on a yearly basis. In Figure 1.2b, I allow the fund to change fees in order to match its expected returns. The flows in Figure 1.2b thus represent changes in the actively managed assets. Even flows of actively managed capital increase more than linearly with cumulative performance.

Economic Motivation. Why should the slope of the flow-performance relationship increase after good performance? In general, the result relies on two properties of the model: (i) Assets under management, K_t, are positively related to the promised value of the manager, $\hat{v}_t$; (ii) the growth rate of the promised value becomes more volatile when ϕ and y increase.

Thanks to the functional assumptions in the model, the relation between assets under management and promised value takes a particularly simple form, $K_t = k_t \hat{v}_t$. However, this relation can be more general. We need two assumptions to obtain a positive relation between assets under management and promised value. First, we need to assume that assets under management enter in the incentive-compatibility constraint and that, to increase assets, the advisor needs to expose the manager to more risk. Second, we need to assume the manager has decreasing absolute risk aversion.

According to the first assumption, if the advisor wants to increase assets, she also has to increase the risk exposure of a risk-averse manager. According to the second

45

assumption, if the manager is promised larger future consumption, he will be more risk tolerant. As a result, the advisor optimally offers more risk and capital to a manager with higher promised value and, hence, higher risk tolerance. This result justifies why assets under management, K_t, are positively related to the promised value of the manager, $\hat{v}_t$. An analogous result holds in the dynamic contracting models in Biais et al. (2010) and DeMarzo and Fishman (2007). These papers show that investments (and disinvestments) at firm level depend on the agent's promised value. Similar to my model, when the agent has a larger promised value, the principal can easily incentivize him to exert more effort and manage a larger firm.

Given this relation between assets and promised value, the percentage change in assets under management for a 1% return is related to the percentage change in the manager's promised value for a 1% return, which coincides with $\hat{\beta}/\sigma$ (see equation (1.18)). I then need to discuss why the growth rate of the promised value becomes more volatile when ϕ and y increase. To understand why, let us fix a level of risk tolerance for the manager by fixing a promised value $\hat{v}$. In deciding how much risk $\hat{\beta}$ to assign to the manager, the advisor faces a trade-off. The advisor would like to increase the manager's risk $\hat{\beta}$ in order to increase capital k and fee revenues. However, doing so is costly for two reasons. First, the advisor has to compensate the risk-averse manager for the additional risk by promising more future consumption. Second, the advisor is committed to low information rents through the multiplier y which penalizes the provision of strong incentives to the manager.

When beliefs ϕ and multiplier y increase, the trade-off tilts in favor of a higher volatility of the manager's promised value, $\hat{\beta}$. When beliefs ϕ are higher, expected

returns and fees, $\mu(\eta, \phi, 0)$, are also higher, so that the benefits of size k and incentives $\hat{\beta}$ increase. As for the cost of the manager's incentives, they decline through two channels. First, because the manager is expected to be more productive (beliefs ϕ increase) and the advisor less constrained (the multiplier y increases), the advisor faces lower costs in delivering future consumption promises. Therefore, the advisor can more cheaply compensate the manager for any additional risk that he has to bear. Second, the commitment of the advisor to low information rents is less binding (the multiplier y increases). Because the benefits of size increase and the cost of incentives decrease, when ϕ and y increase the advisor increases the volatility of the manager's promised value, $\hat{\beta}$, together with k. This result illustrated in Figure 1.1b.

The mechanism that I have discussed so far highlights the dynamic connection between capital flows and managerial incentives. Whereas learning links past performance to expected returns in a standard way, the optimal contract links the slope of the flow-performance relationship to expected returns because of a dynamic trade-off between the costs and the benefits of managerial incentives.

1.5.3 Pay and Performance

Similar to capital, I define the pay-performance sensitivity as the percentage change in compensation for a 1% increase in returns, that is,

$$\epsilon_{\tilde{C}} = \frac{d\tilde{C}_t/\tilde{C}_t}{dR_t}.$$

(a) Pay-performance sensitivity (b) Compensation relative to promised value
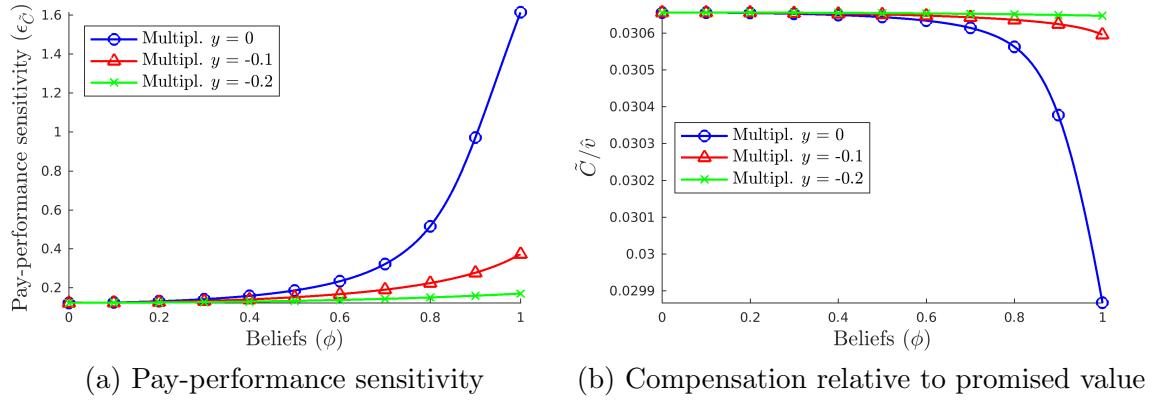
Figure 1.3: Pay-performance sensitivity and compensation in the optimal contract. The parameters of the model are as in Table 1.1.

Because in the optimal contract, compensation takes the form $\tilde{C}_t = \hat{v}_t \tilde{c}(y_t, \phi_t)$ for an optimal compensation-to-value ratio $\tilde{c}(y, \phi) = c(y, \phi)/q(\eta(y, \phi))$, pay-performance sensitivity takes the form

$$
\epsilon_{\tilde{C}}(y, \phi) = \frac{1}{\sigma} \left(\frac{\sigma_{\tilde{c}}(y, \phi)}{\tilde{c}(y, \phi)} + \hat{\beta}(y, \phi) \right),
$$

where $\sigma_{\tilde{c}}(y, \phi)$ is the volatility of $\tilde{c}(y, \phi)$.

From Figure 1.3a, we can observe that the contract implies a bonus for good performance. Because the value of the pay-performance sensitivity, $\epsilon_{\tilde{C}}$, is always positive, compensation increases when the manager realizes good returns. Moreover, because the pay-performance sensitivity $\epsilon_{\tilde{C}}$ is increasing in both beliefs ϕ and the multiplier y, compensation will appear convex in cumulative performance. This result is consistent with the widespread use of convex compensation schemes in the money management industry. For example, Ma et al. (2019) document that mutual fund

48

managers' compensation is often composed of a base salary plus a bonus for good performance. The mechanism driving the convexity of cumulative compensation with respect to cumulative performance is identical to the one behind the convexity of cumulative capital flows.

In addition to the pay-performance sensitivity, we can study whether performance-based compensation is back-loaded or front-loaded, by looking at how current compensation changes relative to future promises, that is, by looking at how $C/\hat{v}$ changes with performance. If $C/\hat{v}$ decreases with beliefs ϕ and multiplier y, we say the performance-based compensation is back-loaded. In this case, after good performance, the advisor increases future promised consumption more than current compensation, thus deferring compensation to the future. Similarly, if $C/\hat{v}$ increases with beliefs ϕ and multiplier y, we say the performance-based compensation is front-loaded.

In the optimal contract, the advisor back-loads the performance-based compensation of the manager, as shown in Figure 1.3b. This result is consistent with common practices in the mutual fund industry that effectively postpone the delivery of compensation to the manager. For example, Ma et al. (2019) show that more than 30% of mutual fund managers are subject to deferred compensation schemes. At the same time, they illustrate that the vast majority of the pay-for-performance schemes rely on the average return over multiple years (on average, three years) in order to determine the bonus. This latest feature effectively implies that, following good performance, the manager can expect an increase in compensation for the next few years, thus effectively back-loading his compensation.

Economic Motivation. Due to the agency frictions, the risk-averse manager faces a risky compensation scheme that is based on his performance. After good performance, the advisor rewards the manager with higher compensation, whereas after bad performance, the manager is punished with lower compensation. However, the riskiness of the compensation must be limited in an optimal contract, because the manager is risk averse. With the exception of highly productive managers, who are subject to very steep incentives, managers face a compensation path that is smoother than returns. This result is consistent with the low point estimates of the pay-performance sensitivity in Ibert et al. (2018).

Similar to the slope of the flow-performance relationship, the pay-performance sensitivity increases after good performance. The motivation is analogous to the one that I have presented for the slope of the flow-performance relationship. Compensation is related to the manager's promised value. This time, however, this relation is based on the manager's desire to smooth consumption. After good performance, the advisor expects higher returns from the manager, and therefore desires to allocate more capital to the manager. However, to ensure incentive compatibility with a larger amount of assets under management, the advisor needs to expose the manager to a riskier promised value. Given that compensation is related to promised value, a riskier promised value for the manager translates into riskier compensation.

To understand why the advisor back-loads the manager's performance-based compensation, let us consider the trade-off that compensation involves. To study this trade-off, fix a level of experimentation η and, hence, a level of cost $q(\eta)$, so that we can think of compensation and consumption as proportional. By decreasing the

consumption-to-promised-value ratio, c, the advisor increases the growth rate of the manager's promised value (see (1.18)). A higher promised value involves a trade-off. On the one hand, the advisor faces higher costs from a larger promised value, because she has to deliver larger future compensation. On the other hand, the advisor benefits from a larger promised value, because if the manager has a large promised utility, he can tolerate high levels of risk. Therefore, the advisor can exploit this higher risk tolerance to expose the manager to more risk and increase assets under management and fee revenues.

After a good return, the trade-off tilts in favor of a larger continuation value and thus of a lower consumption-to-promised-value ratio c. On the one hand, the advisor can now deliver any given promised utility at a lower cost because, after a good shock, the manager is expected to be more productive (beliefs ϕ increase) and the commitment of the advisor to low information rents is less binding (multiplier y increases). On the other hand, because the advisor expects higher returns from the manager, she has stronger incentives to increase the manager's assets under management. Hence, the advisor desires to increase the risk tolerance of the manager through a higher promised value. Consequently, after good performance, the advisor reduces the ratio of consumption to promised utility in order increase the growth rate of the promised value. By doing so, the advisor back-loads the performance-based compensation of the manager.

(a) Drift of multiplier y (b) Incentives
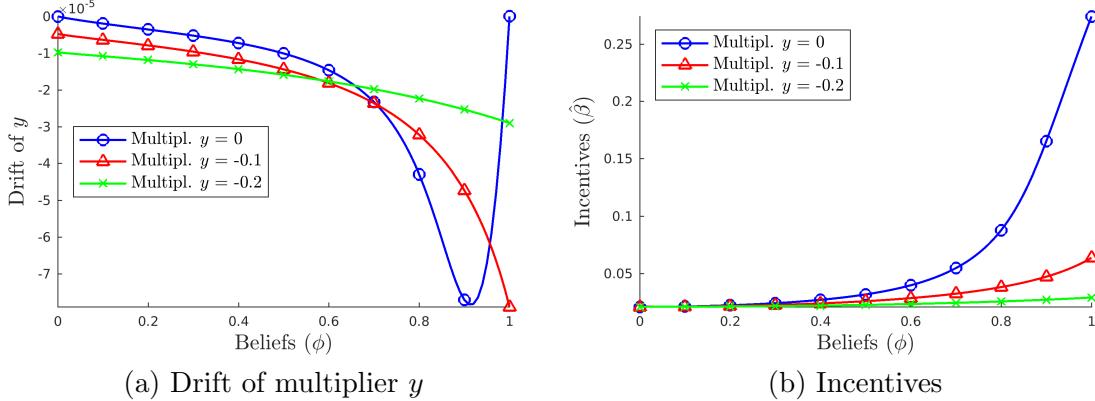
Figure 1.4: Drift of multiplier y and incentives in the optimal contract. The parameters of the model are as in Table 1.1.

1.5.4   Long-Term Implications and the Dynamics of Multiplier y

A full-commitment contract bears implications for flows, incentives, and performance in the long run. While beliefs ϕ are martingales, the multiplier y drifts down over time, as shown in Figure 1.4a. The negative drift of y is a robust outcome that I find across all the parameterizations that I have explored. To understand the implications of a negative drift of y in terms of flows, incentives, and performance, recall that we can interpret $-y$ as a penalty for incentives and learning. If the multiplier y becomes more negative, the advisor faces a large penalty for incentives, $\hat{\beta}$, and for learning, η. Consequently, by giving a negative drift to y, the advisor is committing herself to reduce incentives and expected performance over time.

By reducing incentives $\hat{\beta}$ over time, the advisor reduces the slope of the flow-performance relationship and the pay-performance sensitivity. Figures 1.1a and 1.3a show that for increasingly more negative values of the multiplier y, flows and pay react increasingly less to performance. Moreover, for very negative values of the

(a) Experimentation   (b) Expected excess return
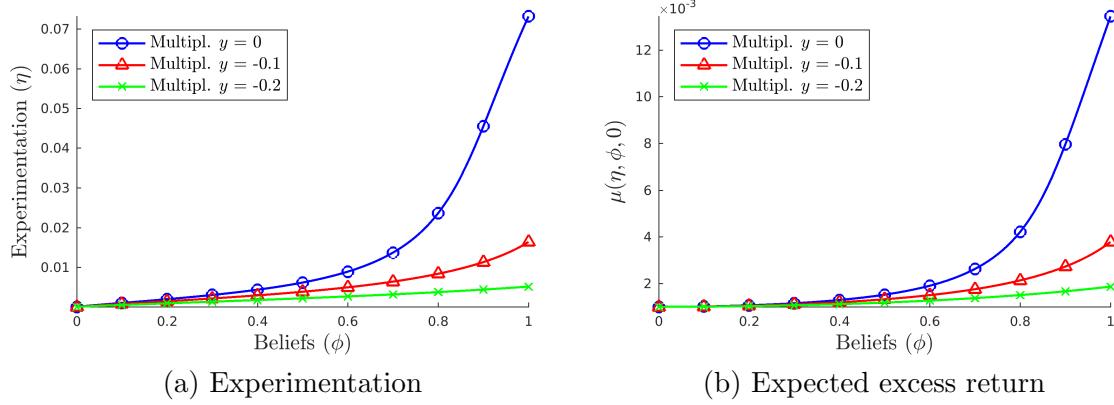
Figure 1.5: Experimentation and performance in the optimal contract. The parameters of the model are as in Table 1.1.

multiplier y, the slope of the flow-performance relationship is barely affected by past performance, thus suggesting the history dependence of the slope is, on average, substantially smaller for managers with long tenure. I test this prediction in section 1.6.3.

By reducing experimentation η over time, the advisor also reduces the productivity of the manager. In Figure 1.5, we see that for very negative values of the multiplier y, the advisor requires low experimentation η from the manager, which results in lower expected returns. Almazan et al. (2004) show that older fund managers are subject to more investment constraints, which, effectively, may limit the informativeness of their returns. As long as age and tenure are correlated, my model could provide one additional framework to interpret these empirical findings. Consistently with an optimal contract, more experienced managers may simply be required to undertake less sophisticated investments and to be subject to lower risk.

Economic Motivation. To understand why the advisor wants to reduce the incentives and the experimentation of the manager over time, let us consider the incentive constraint (1.8) and the manager's information rent. The advisor faces a trade-off between ex ante and ex post efficiency. From an ex ante perspective, the advisor wants to minimize the manager's information rent. The information rent is costly, because it represents a risk that the advisor cannot exploit to provide incentives to the risk-averse manager. However, from equation (1.11), the manager's information rent corresponds to the present value of future incentives and experimentation. If the advisor designs a contract with a small information rent, she has to commit to future experimentation η and incentives $\hat{\beta}$ that are inefficiently low from an ex post perspective. The commitment of the advisor to these ex post inefficient incentives and experimentation is captured by the multiplier y. By giving a negative drift to the multiplier, the advisor increases the penalty for incentives and experimentation over time. To sum up, in designing the ex ante optimal contract, the advisor commits to decreasing incentives and learning over time through the negative drift of multiplier y. By taking ex post inefficient actions, she can reduce the ex ante information rent of the manager, and thus relax the incentive-compatibility constraint (1.8).

Full commitment by the advisor is crucial for this result. After any period of time, the advisor is tempted to renegotiate the contract, leaving the manager with an unchanged continuation value, but increasing his information rent to the point at which the multiplier y is equal to 0. In Appendix 1.A, I study contracts in which advisor does not commit to an information rent. As expected, incentives $\hat{\beta}$ and experimentation η will simply be a function of beliefs ϕ, because no other state variable

enforces ex post inefficient choices of incentives and experimentation. However, all the other predictions of the model about the flow-performance relationship and the manager's compensation continue to robustly hold, regardless on the commitment of the advisor.

In Appendix 1.B, I provide a simple two-period contracting model with learning. This model provides an additional example of the trade-off between ex ante and ex post optimality that the principal faces. This two-period model is substantially different from the dynamic model of the paper. However, I can analytically show that the advisor commits to low experimentation ex post in order to achieve ex ante optimality. The two-period model hence provides further support for the results of this section and allows us to frame these results as general outcomes of contracting models with learning.

1.6   Empirical Tests

As shown in the previous section, the model generates a slope of the flow-performance relationship that varies over time in response to the manager's past performance. The model highlights the dynamic nature of the flow-performance relationship and its connection to the manager's incentives.

In this section, I test the predictions of the model. From the discussion in section 1.5.2, we should empirically observe that, everything else being equal, funds with better past performance display a steeper slope in their flow-performance relationship. Moreover, based on section 1.5.4, we should observe that past performance has a weaker effect on the steepness of the flow-performance relationship if the manager

has longer tenure.

I test the model's prediction in mutual fund data. I focus on US mutual funds investing in US equity, which provide a sample of money managers with very liquid and ample investment opportunities. In principle, the general economic mechanism of the model could apply to other money managers such as private equity funds, hedge funds, and bond funds. However, previous studies point out that these intermediaries face capacity constraints or undertake illiquid investments. These frictions on the assets side of the intermediary may mechanically introduce concavity in the flow-performance relationship.[11] Because my model, for tractability reasons, abstracts from size constraints and illiquidity, the sample of US equity mutual fund constitutes the natural testing ground of my theory.

1.6.1 Data and Variables of Interest

The sample of mutual funds comes from the Center for Research in Security Prices (CRSP). I consider monthly observations from December 1999 to December 2018.

I include only actively managed funds that invest in US stocks, thus excluding index funds, ETFs, and bond and commodity funds. Because Elton et al. (2001) document an upward bias in the performance of small funds, I include funds starting

---

11. Kaplan and Schoar (2005) find that the relation between flows and performance is concave in private equity funds. They explain this result by observing that private equity funds may have access to a limited number of deals, and that the human capital of general partners cannot be easily scaled. Getmansky et al. (2004) illustrate that hedge funds' returns are highly serially correlated, and they show that a likely cause for the serial correlation is the hedge funds' exposure to illiquid securities. Goldstein et al. (2017) show that the relation between flows and performance is concave in corporate bond mutual funds. As a possible explanation, they suggest that because of the illiquidity of the corporate bond market, investors face strategic complementarities when fleeing a fund with bad performance (Chen et al., 2010).

from the date at which their assets under management exceed 15 million in 2011 dollars, as in Pastor et al. (2015). Finally, I restrict the sample to funds that are open to new investors.[12]

CRSP reports net monthly returns and annual expense ratios for every share class. I compute the gross returns of every share class by adding 1/12 of the expense ratio to the net monthly return. I then compute fund-level gross returns and expense ratios by taking a weighted average of these quantities across each fund's share classes, where the weights are given by the total net asset value of each share class. Pastor et al. (2015) observe that actively managed funds are unlikely to charge less than a 0.1% annual fee. Thus, following their procedure, I exclude observations whose fees are below the 0.1% threshold, since these observations might represent index funds or data entry mistakes. I also exclude observations whenever fees exceed 10% per year.

The model predicts that the manager's tenure matters for the dynamic behavior of the flow-performance relationship. To test this prediction, I compute the tenure of the manager at the fund management company. CRSP provides the tenure of managers at each fund, which is not the relevant measure of tenure according to my model. However, I can use this information, together with managers' and companies' identifiers, to construct a measure of the length of the contractual relationship between managers and fund companies.[13] Funds may have multiple managers. When I

---

12. By excluding small funds and funds closed to new investors, I avoid the incubation bias identified by Evans (2010). Moreover, in my empirical analysis I control for fund flows over the previous 12 months, thus effectively excluding funds of less than one year of age.

13. For each manager-company pair, I define the manager start date as the date at which the manager first started working for the management company. I obtain this date by looking at when

study the relation between tenure and the flow-performance relationship, I consider the tenure of the most senior manager.

The model yields predictions about the relation between flows of capital, current performance, and past performance. I measure the net flow of capital into fund i at time $t + 1$ as

$$F_{it+1} = \frac{K_{it+1} - K_{it}}{K_{it}} - R^N_{it+1},$$ (1.22)

where K_{it} are the assets under management of fund i at time t, and R^N_{it+1} is the net return that fund i delivers to investors from time t to time $t + 1$.

To measure performance, first I compute the fund's benchmark return as the average gross return of funds with the same investment objective of the fund. Hunter et al. (2014) show this benchmark return accounts for commonalities across similar active strategies. Then I obtain the fund's performance as the gross return of the fund in excess of its style benchmark, that is,

$$\tilde{R}_{it} = R_{it} - \bar{R}_{s(i)t},$$ (1.23)

where R_{it} is the gross return of fund i from $t - 1$ to t, $s(i)$ is the style of fund i (as described by the CRSP objective classification), and $\bar{R}_{st}$ is the average gross return of all funds with the same investment objective s. Khorana et al. (2007) and Spiegel and Zhang (2013) use analogous measures of funds' objective-adjusted performance.

---

the manager-company pair first appears in the dataset and at CRSP-reported start dates of the manager in any of the company's funds. This algorithm allows me to keep track of managers within a management company. However, it does not allow me to track transfers of managers across companies. After this procedure, I can compute the tenure of the manager in a management company.

I measure the past performance of a fund manager by considering the average performance of the fund over the previous l months (provided that the manager of the fund did not change in those l months),

$$\text{PastPerf}_{i[t-l,t-1]} = \frac{1}{l} \sum_{j=1}^{l} \tilde{R}_{it-j}.$$ (1.24)

Finally, to improve clarity in my subsequent discussion, I introduce a separate notation for the cumulative performance of the manager,

$$\text{CumPerf}_{i[t-l,t]} = \text{PastPerf}_{i[t-l,t]}.$$

I refer to cumulative performance $\text{CumPerf}_{i[t-l,t]}$ at month t when I consider a measure of performance up to and including the current month t, whereas I refer to past performance $\text{PastPerf}_{i[t-l,t-1]}$ at month t when I consider a measure of performance that does not include month t.

Berk and van Binsbergen (2015) and Pastor et al. (2015) document that the CRSP database contain data-entry mistakes, some of them representing large outliers in the distribution of flows and returns. To avoid the risk that my results are driven by outliers, I remove the tails of the distributions of capital flows and gross returns, keeping only observations within the 1st and 99th percentiles.

In Table 1.2, I report the summary statistics of the data used in my empirical analysis. The final sample contains 3,903 funds, 7,635 fund-manager observations, and 229 months. Table 1.3 contains the description of the variables used in the empirical analysis. In this section, I compute past performance over the previous 6

Table 1.2: Summary statistics. The sample contains 3,903 funds, 7,635 fund-manager pairs, and 229 months.

Variable	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
F_{it+1}	283,111	−0.001	0.038	−0.173	−0.015	−0.005	0.007	0.304
$\tilde{R}_{it}$	283,111	0.00004	0.019	−0.237	−0.009	0.000	0.009	0.193
Tenure$_{it}$ (years)	283,111	11.671	8.292	0.085	5.490	9.753	16.005	62.038
K_{it} (USD mln)	283,111	1,209	3,947	0.1	61	228	880	135,373
ExpRatio$_{it}$ (% per year)	282,966	1.235	0.498	0.100	0.950	1.180	1.455	9.950
FundAge$_{it}$ (years)	283,095	15.493	13.576	0.493	6.499	12.008	19.553	94.526

Table 1.3: Definition of variables

Variable	Definition	Description
F_{it+1}	Net flow	Growth rate of AUM from month t to month $t+1$ minus net return
$\tilde{R}_{it}$	(Current) performance	Gross return of fund i in excess of its style benchmark
$\bar{R}_{st}$	Style benchmark	Equally weighted average gross return of all funds with style s
PastPerf$_{i[t-6,t-1]}$	Past performance	Manager's average performance from month $t-l$ to month $t-1$
CumPerf$_{i[t-l,t]}$	Cumulative performance	Manager's average performance from month $t-l$ to month t
X_{it}	Controls	12 lags of net flows into the fund, log of fund size, expense ratio, log of fund age, and log of the manager's tenure
ι_{it}^{FMgr}	Fund-manager fixed effect	Fixed effect for each fund-manager pair
ι_{it}^{SMon}	Style-month fixed effect	Monthly fixed effects for funds with the same investment objective

months, that is $l = 6$. In Appendix 1.F, I check the robustness of the results by considering past performance over the previous 12 months, that is, $l = 12$.

1.6.2 Prediction 1: History Dependence of the Flow-Performance Relationship

The first prediction of the model can be summarized as follows.

Prediction 1. The slope of the monthly flow-performance relationship increases with past performance.

As a baseline test, I estimate the regression equation

$$
F_{it+1} = a_1 \tilde{R}_{it} + a_2 \text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it} + a_3 \text{PastPerf}_{i[t-l,t-1]}
$$

$$
+ a_4 (\text{PastPerf}_{i[t-l,t-1]})^2 \tag{1.25}
$$

$$
+ a_5 \tilde{R}_{it}^2 + c' X_{it} + \iota_{it}^{FMgr} + \iota_{it}^{SMon} + u_t,
$$

and test whether $a_2 > 0$. A positive coefficient on the term that capture the inter-action between returns and past performance, $\text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it}$, indicates that the flow-performance slope is an increasing function of past performance. Whereas the model predicts that a_1 is also positive, testing for $a_1 > 0$ alone cannot be taken as evidence for my theory. Other theories, for example, Berk and Green (2004) and Lynch and Musto (2003), predict a positive coefficient a_1. Controls X_{it} include 12 lags of monthly net flows into the fund, the logarithm of fund size, its expense ratio, the logarithm of the fund's age, and the logarithm of the manager's tenure. ι_{it}^{FMgr} is a fixed effect for the fund-manager pair and ι_{it}^{SMon} is a style-month fixed effect. Standard errors are double-clustered at the month and at the fund level. Because I measure past performance at the fund-manager level, for an observation to be in-cluded in the estimation we need a history of at least $l + 1$ months of performance for each fund-manager pair.

Results. Columns (1) and (2) of Table 1.4 provide estimates of model (1.25), where the past performance of the manager is computed over the previous six months, that is $l = 6$.

The data support the model's prediction that the slope of the flow-performance

Table 1.4: Effect of past performance on the slope of the flow-performance relationship. Past performance is computed over 6 months. (Continues.)

	F_{it+1} (Net Flow)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{R}_{it}$	0.172***	0.166***	0.172***	0.166***	0.135***	0.130***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)
$\text{PastPerf}_{i,[t-6,t-1]} \cdot \tilde{R}_{it}$	2.562***	2.270***	2.374***	2.134***		
	(0.670)	(0.678)	(0.625)	(0.609)		
$\tilde{R}_{it} \cdot \mathbb{I}[\text{PastPerf}_{i,[t-6,t-1]} > 0]$					0.067***	0.066***
					(0.015)	(0.015)
$\text{PastPerf}_{i,[t-6,t-1]} \cdot \mathbb{I}[\tilde{R}_{it} > 0]$					0.052*	0.054*
					(0.031)	(0.032)
$\text{PastPerf}_{i,[t-6,t-1]}$	0.354***	0.399***	0.354***	0.399***	0.244***	0.289***
	(0.027)	(0.026)	(0.027)	(0.026)	(0.039)	(0.038)
$(\text{PastPerf}_{i,[t-6,t-1]})^2$	0.566	0.408			0.316	1.137
	(0.944)	(0.960)			(2.572)	(2.624)
$(\text{CumPerf}_{i,[t-6,t]})^2$			0.770	0.555	0.666	−0.596
			(1.285)	(1.306)	(3.742)	(3.905)
$\mathbb{I}[\tilde{R}_{it} > 0]$					0.000	0.000
					(0.000)	(0.000)
$\mathbb{I}[\text{PastPerf}_{i,[t-6,t-1]} > 0]$					0.002***	0.002***
					(0.000)	(0.000)
$\tilde{R}_{it}^2$	0.571***	0.621***	0.555***	0.610***	0.558***	0.645***
	(0.128)	(0.136)	(0.130)	(0.135)	(0.140)	(0.147)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Style-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Manager FE	No	Yes	No	Yes	No	Yes
Observations	186,192	186,192	186,192	186,192	186,192	186,192
R^2	0.347	0.378	0.347	0.378	0.348	0.378

Notes: $^{*}p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 1.4: (Continued.) The history dependence of the slope of the flow-performance relationship is measured by the coefficient on $\text{PastPerf}_{i[t-6,t-1]} \cdot \tilde{R}_{it}$. $\tilde{R}_{it}$ measures current performance and is calculated, for each month t, as the gross return of fund i in excess of the equally weighted average gross return of all funds with the same style. $\text{PastPerf}_{i[t-6,t-1]}$ measures the past performance of the manager and is calculated as the average excess return over the style benchmark in the six months from $t-6$ to $t-1$. The dependent variable, F_{it+1}, measures the net flow of capital and is calculated as the growth rate of assets under management from month t to month $t+1$ minus the net return over the same period. $\text{CumPerf}_{i[t-6,t]}$ is the average performance of the manager over the style benchmark in the months from $t-6$ to t. $\mathbb{I}[\cdot]$ is the indicator function. Controls include 12 lags of monthly net flows into the fund, the log of fund size, its expense ratio, the log of fund age, and the log of the manager's tenure. Standard errors are in parentheses and they are double-clustered at the fund and at the month level.

relationship positively depends on the history of performance, as reflected by the positive coefficients on the $\text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$ term. One may suspect that a positive coefficient on $\text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$ captures the convexity of flows in the cumulative performance up to month t, that is, convexity in $\text{CumPerf}_{i[t-6,t]}$. In columns (3) and (4), I replace the square of past performance with the square of cumulative performance up to and including month t. Because of the correlation between $\text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$ and $(\text{CumPerf}_{i[t-l,t]})^2$, the coefficient on the interaction term $\text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$ decreases. However, it remains positive and statistically significant. This result highlights that the flow-performance slope increases with past performance even after controlling for the convexity of flows with respect to cumulative performance up to the current month.

In columns (5) and (6), I address two possible concerns. First, one may suspect that a positive coefficient on the interaction term $\text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$ captures a convexity in cumulative performance that the quadratic term $(\text{CumPerf}_{i[t-l,t]})^2$ does not fully capture. Second, one may suspect that a positive coefficient on
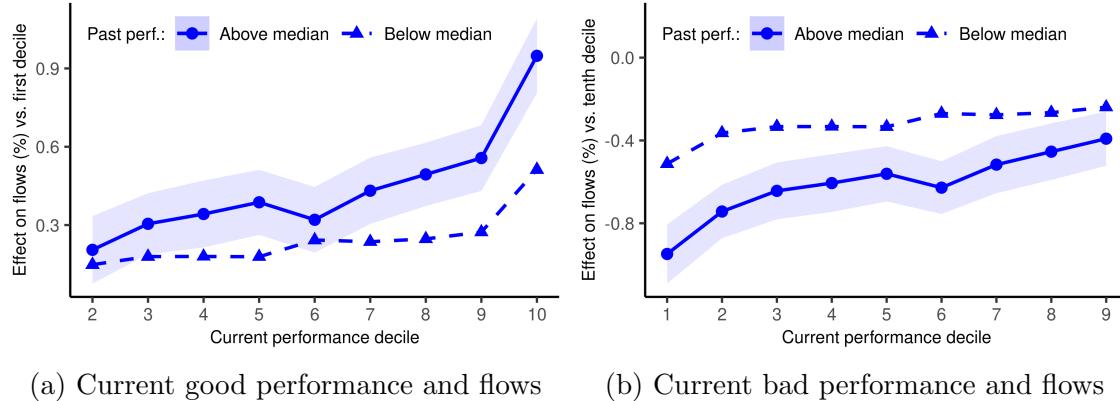
(a) Current good performance and flows   (b) Current bad performance and flows

Figure 1.6: History dependence of the relation between flows and performance. Past performance is computed over 6 months. Figures (a) and (b) show how flows change with current performance and how the change depends on past performance. I sort funds into deciles based on their current performance and into halves based on their past performance. Past performance is the average excess return over the style benchmark in the previous 6 months. I then run regression of flows on dummies for the deciles of current performance, dummies for the halves of past performance, and interactions between the two sets of dummies. As controls, I include dummy variables for cumulative performance $\text{CumPerf}_{i[t-6,t]}$ sorted into deciles, 12 lagged flows, the logarithm of fund age, the logarithm of the manager's tenure, the logarithm of lagged assets under management, fund fees, fund-manager fixed effects, and style-month fixed effects. The shaded areas represent 95% confidence intervals for the change in the effect of current performance on flows when past performance increases above the median. Confidence intervals are constructed by double-clustering standard errors at the month and at the fund level.

In Figure (a), I plot the effect of current good performance (that is, performance relative to the first decile) on flows, while, in Figure (b), I plot the effect of current bad performance (that is, performance relative to the tenth decile) on flows.

$\text{PastPerf}_{i[t-6,t-1]}\tilde{R}_{it}$ actually indicates that flows become more sensitive to past performance after good current performance. In columns (5) and (6), I show that my results are not affected by these concerns: When current performance is positive, there is only weak statistical evidence that flows are more sensitive to past performance. On the contrary, when past performance is positive, there is strong

statistical evidence that flows become more sensitive to current performance. These results support the prediction of the model that good past performance increases the sensitivity of flows to current performance.

In Figure 1.6, I use a non-parametric approach to show that flows become more sensitive to current performance after better performance in the past. In Figure 1.6a, for every month I sort past performance $\text{PastPerf}_{i[t-6,t-1]}$ into halves (above and below median), and I sort current performance $\tilde{R}_{it}$ into deciles. I then regress flows on dummies for the deciles of current performance, dummies for the halves of past performance, and interactions between the two sets of dummies. I also sort cumulative performance $\text{CumPerf}_{i[t-6,t]}$ into deciles and add dummy variables for cumulative performance deciles as controls. I include the same controls X_{it} of equation (1.25), fund-manager fixed effects, and style-month fixed effects. Standard errors are double-clustered at the month and at the fund level.

Figure 1.6a plots the incremental effect of current performance on flows (relative to the first decile of current performance), and it shows that such incremental effect is larger for above-median past performers. For example, the blue dot at decile 10 represents the increase in flows for a fund with top current performance, relative to a fund with bad current performance, conditional on past performance being above the median. The shaded blue area represents the 95% confidence interval for the incremental effect of past performance. If the effect of current performance for below-median past performers (the triangle) lies outside the shaded area, then we reject, at a 95% confidence level, the hypothesis that the relation between flows and current performance is the same for above-median past performers (the dot) and

65

below-median past performers (the triangle).

Someone may suspect that the history-dependence of the flow-performance relationship is due to investors' inertia. For example, investors may wait to observe good performance both in the past and in the current month before investing in a fund. This hypothesis would imply that, after good performance in the past, flows are less sensitive to bad performance. Figure 1.6b illustrates the opposite happens in the data. The figure plots the effect of current bad performance (that is, performance relative to the tenth decile of current performance) on flows, and it highlights that flows are more sensitive to bad performance for funds that experienced above-median performance in the past. Consistently with my model, after good performance in the past, flows are more sensitive not only to current good performance, but also to current bad performance. We can therefore reject the hypothesis that the history-dependence of the flow-performance relationship is due to investors' inertia.

1.6.3   Prediction 2: Flows, Performance, and Managers' Tenure

The model predicts a relation between the manager's tenure and the history dependence of the flow-performance slope. The prediction can be stated as follows.

Prediction 2.   The slope of the flow-performance relationship increases less with past performance if the manager has longer tenure.

Two mechanisms drive this prediction. The first mechanism is the optimal contract itself. As discussed in section 1.5, in the long run, the advisor will constrains

66

the manager's experimentation and incentives up to the point where the slope of the flow-performance relationship will no longer change with past performance. The second mechanism is the convergence of beliefs. If the manager's skill is a fixed (unknown) parameter, then, in the long run, beliefs will converge in probability to the true value, and a negligible amount of learning will take place.

I test the second prediction of the model by using the following regression:

$$
\begin{aligned}
F_{it+1} = \quad & a_0 + a_1 \tilde{R}_{it} + a_2 \text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it} \\
& + a_3 \text{PastPerf}_{i[t-l,t-1]} + a_4 (\text{PastPerf}_{i[t-l,t]})^2 + a_5 \tilde{R}_{it}^2 \\
& \sum_{j=2}^{5} \text{TenureQuintile}_{it}^{j} \left(a_0^{Tj} + a_1^{Tj} \tilde{R}_{it} + a_2^{Tj} \text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it} \right. \\
& \left. + a_3^{Tj} \text{PastPerf}_{i[t-l,t-1]} + a_4^{Tj} (\text{PastPerf}_{i[t-l,t]})^2 + a_5^{Tj} \tilde{R}_{it}^2 \right) \\
& \sum_{j=2}^{5} \text{AgeQuintile}_{it}^{j} \left(a_0^{Aj} + a_1^{Aj} \tilde{R}_{it} + a_2^{Aj} \text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it} \right. \\
& \left. + a_3^{Aj} \text{PastPerf}_{i[t-l,t-1]} + a_4^{Aj} (\text{PastPerf}_{i[t-l,t]})^2 + a_5^{Aj} \tilde{R}_{it}^2 \right) \\
& \sum_{j=2}^{5} \text{SizeQuintile}_{it}^{j} \left(a_0^{Sj} + a_1^{Sj} \tilde{R}_{it} + a_2^{Sj} \text{PastPerf}_{i[t-l,t-1]} \tilde{R}_{it} \right. \\
& \left. + a_3^{Sj} \text{PastPerf}_{i[t-l,t-1]} + a_4^{Sj} (\text{PastPerf}_{i[t-l,t]})^2 + a_5^{Sj} \tilde{R}_{it}^2 \right) \\
& + c' X_{it} + \iota_{it}^{FMgr} + \iota_{it}^{SMon} + u_t.
\end{aligned}
$$

(1.26)

$\text{TenureQuintile}_{it}^{j} = 1$ if, in month t, the tenure of the manager of fund i belongs to the jth quintile of the distribution of managerial tenure in month t, otherwise $\text{TenureQuintile}_{it}^{j} = 0$. Similarly, $\text{AgeQuintile}_{it}^{j} = 1$ if, in month t, the age of fund

(a) Effect of managerial tenure
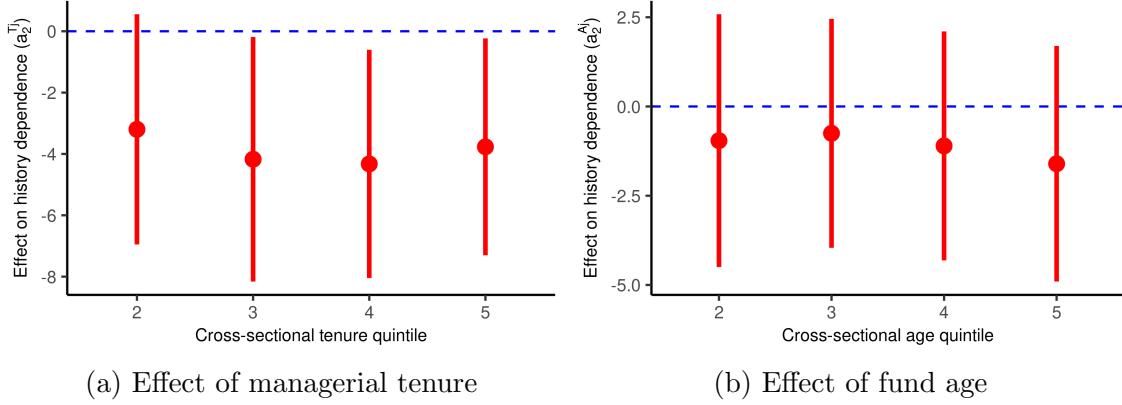
(b) Effect of fund age

Figure 1.7: Effect of managerial tenure and fund age on the history dependence of the relation between flows and current performance. Past performance is computed over 6 months. (Continues.)

i belongs to the jth quintile of the distribution of fund age in month t, otherwise $\text{AgeQuintile}_{it}^j = 0$. $\text{SizeQuintile}_{it}^j = 1$ if, in month t, the size of fund i belongs to the jth quintile of the distribution of fund size in month t, otherwise $\text{SizeQuintile}_{it}^j = 0$.

I verify Prediction 2 by testing whether $a_2^{Tj} < 0$ for $j = 2, \ldots, 5$. As in regression (1.25), controls X_{it} include 12 lags of monthly net flows into the fund, the logarithm of fund size, its expense ratio, the logarithm of fund age, and the logarithm of the manager's tenure. I also control for fund-manager fixed effects, ι_{it}^{FMgr}, and style-month fixed effects, ι_{it}^{SMon}. Standard errors are double-clustered at the month and at the fund level.

Results. Figure 1.7a shows that slope of the flow-performance relationship increases less with past performance if the manager has longer tenure.The dots at quintiles 2 to 5 are the estimates of a_2^{Tj} for $j = 2, \ldots, 5$, and the vertical lines represent 90% confidence intervals for the incremental effect of managerial tenure.

68

Figure 1.7: (Continued.) I run regression

$$F_{it+1} = \quad a_0 + a_1 \tilde{R}_{it} + a_2 \text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it}$$
$$+ a_3 \text{PastPerf}_{i[t-6,t-1]} + a_4 (\text{PastPerf}_{i[t-6,t]})^2 + a_5 \tilde{R}_{it}^2$$

$$\sum_{j=2}^{5} \text{TenureQuintile}_{it}^j \left(a_0^{Tj} + a_1^{Tj} \tilde{R}_{it} + a_2^{Tj} \text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it} \right.$$
$$\left. + a_3^{Tj} \text{PastPerf}_{i[t-6,t-1]} + a_4^{Tj} (\text{PastPerf}_{i[t-6,t]})^2 + a_5^{Tj} \tilde{R}_{it}^2 \right)$$

$$\sum_{j=2}^{5} \text{AgeQuintile}_{it}^j \left(a_0^{Aj} + a_1^{Aj} \tilde{R}_{it} + a_2^{Aj} \text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it} \right.$$
$$\left. + a_3^{Aj} \text{PastPerf}_{i[t-6,t-1]} + a_4^{Aj} (\text{PastPerf}_{i[t-6,t]})^2 + a_5^{Aj} \tilde{R}_{it}^2 \right)$$

$$\sum_{j=2}^{5} \text{SizeQuintile}_{it}^j \left(a_0^{Sj} + a_1^{Sj} \tilde{R}_{it} + a_2^{Sj} \text{PastPerf}_{i[t-6,t-1]} \tilde{R}_{it} \right.$$
$$\left. + a_3^{Sj} \text{PastPerf}_{i[t-6,t-1]} + a_4^{Sj} (\text{PastPerf}_{i[t-6,t]})^2 + a_5^{Sj} \tilde{R}_{it}^2 \right)$$

$$+ c' X_{it} + \iota_{it}^{FMgr} + \iota_{it}^{SMon} + u_t,$$

where $\text{TenureQuintile}_{it}^j = 1$ if, in month t, the tenure of the manager of fund i belongs to the jth quintile of the distribution of managerial tenure in month t; $\text{AgeQuintile}_{it}^j = 1$ if, in month t, the age of fund i belongs to the jth quintile of the distribution of fund age in month t; $\text{SizeQuintile}_{it}^j = 1$ if, in month t, the size of fund i belongs to the jth quintile of the distribution of fund size in month t. Table 1.3 contains the description of all the other variables used in the regression. Standard errors are double-clustered at the month and at the fund level.

In Figure (a), the dots in the figure represents estimated coefficients a_2^{Tj}'s. The vertical lines represent 90% confidence intervals for the incremental effect of tenure on the flow-performance slope relative to the first quintile: If the vertical red line at quintile j does not cross the dashed horizontal line, then we reject the hypothesis that $a_2^{Tj} \geq 0$ at a 95% confidence level.

Figure (b) is the analogous of Figure (a) for the effect of fund age on the history-dependence of the flow-performance relationship.

Because the confidence intervals do not contain the zero for quintiles 3 to 5 (graphically, the red vertical lines do not intersect the horizontal dashed line), we reject the hypothesis that $a_2^{Tj} \geq 0$ at a 95% confidence level for $j = 3, \ldots, 5$. In order words,

69

the slope of the flow-performance relationship increases less with past performance for funds with more senior managers.

Figure 1.7b is analogous to Figure 1.7a, but plots the incremental effect of fund age on the history dependence of the flow-performance relationship. No clear correlation exits between the fund's age and the extent to which the flow-performance slope depends on the history of performance. Compared to previous studies, my model and my empirical analysis establish a connection between managerial tenure and the history dependence of the flow-performance slope, rather than between fund age and the value of the slope. Therefore, my results complement, but do not overlap with, the results of Chevalier and Ellison (1997), who show that the flow-performance relationship is less steep in older funds. Similarly, the theoretical predictions of my model are distinct from those of Berk and Green (2004), who explain why older funds have a flatter flow-performance relationship.

1.7   Conclusions

In this paper, I study how mutual fund flows respond to performance when portfolio managers face optimal incentive contracts under moral hazard and learning. I show that both flows and managerial compensation increase more than linearly in cumulative past performance, consistent with empirical evidence.

I develop a dynamic model that explicitly takes into account the two challenges that money management firms face: raising capital from investors and providing incentives to portfolio managers. I illustrate the connection between the two challenges and the optimal strategy that money management firms should undertake. I

70

show that the empirical patterns in capital flows and managerial compensation are consistent with this strategy.

The model highlights the dynamic nature of the relation between flows and performance. In particular, the model offers novel testable predictions on the dynamics of the flow-performance relationship. First, after a history of good performance, flows react more strongly to current performance. Second, the flow-performance relationship depends less on the history of performance for managers with longer tenure. I test these predictions in mutual fund data and provide empirical support for the model.

By focusing on contracting frictions inside the money management industry, the model offers a basis for further research on financial intermediaries. Since I provided a partial equilibrium model, a natural extension would be to explore the asset-pricing implications of optimal money management contracts. Building on this extension, we can ask how monetary policy in the form of quantitative easing affects incentives in the money management industry and to what extent the intermediary sector facilitates or hinders the transmission of monetary policy to the real economy. Finally, we can modify the model and make it suitable to study optimal contracts for other types of money managers. For example, by considering illiquid investments and Poisson shocks, we could explain common patterns in the hedge fund industry, such as the use of lock-up provisions and high-water marks.

Appendix 1.A   Alternative Contractual Environments

The design and implementation of the optimal contract requires the full commitment of the advisor to the initial promises. The ex ante optimal contract implies actions that are ex post inefficient. For example, the advisor would like to renegotiate the contract at any time to reset $Y = 0$. Similarly, a contract with full commitment is not robust to competition between advisors or to the assumption that the manager could leave the advisor to open his own fund. In this section, I explore these two alternative scenarios. In the first one, I consider a renegotiation-proof (or state-contingent) contract. In the second one, I explore the incentives that a market of atomistic investors can provide to a manager without the full commitment of an advisor.

I first present these two contractual environment formally, then show their equivalence and discuss their outcomes.

1.A.1   Renegotiation-Proof Contract

To understand the intuition behind a renegotiation-proof contract, imagine that multiple advisors are competing for the same manager. Suppose that one of them offers the contract in section 1.4 to the manager. After some time has passed, the advisor will be committed to undertaking ex post inefficient actions, captured by a strictly negative multiplier $Y_t < 0$. Let $\hat{\xi}(V_t, Y_t, \phi_t)$ be the information rent implied by this contract at time t. At that point, the advisor is willing to transfer wealth $J(V_t, \hat{\xi}(V_t, Y_t, \phi_t), \phi_t)$ to another advisor who will then employ the man-

ager by promising the same continuation value V_t, but with a re-set information rent $\hat{\xi}(V_t, 0, \phi_t)$. At these terms, the first advisor would be indifferent between keeping the manager and transferring him. The manager is also indifferent, because he obtains the same continuation value with the old and the new advisor. However, the new advisor makes a strictly positive gain. She obtains a wealth transfer of $J(V_t, \hat{\xi}(V_t, Y_t, \phi_t), \phi_t)$ from the first advisor, but she will have to bear costs $J(V_t, \hat{\xi}(V_t, 0, \phi_t), \phi_t)$ to employ the manager, where $J(V_t, \hat{\xi}(V_t, 0, \phi_t), \phi_t) < J(V_t, \hat{\xi}(V_t, Y_t, \phi_t), \phi_t)$.

Therefore, if the manager can be transferred across funds and an advisor cannot commit to retain the manager, the contract of section 1.4 would not be credible. The terms of the contract would be continuously renegotiated, making it impossible to implement an ex ante optimal contract that requires ex post inefficient choices.

I therefore develop a notion of renegotiation-proof contracts that offer a credible alternative to optimal contracts when the advisor cannot fully commit to the terms of the contract. The advisor may lack commitment because the manager can be transferred across funds, as discussed above. Alternatively, the advisor may lack commitment because, after writing an initial contract, the advisor and the manager could mutually agree to change the terms of the contract ex post, as in the "commitment and renegotiation" model of Laffont and Tirole (1990).

Define the following set:

$$\mathcal{I}(V, \phi, t) = \left\{ \mathcal{C}_t : \mathcal{C}_t \text{is IC}, \mathrm{E}\left[\int_t^\infty e^{-R(s-t)} u(C_s + m_s \lambda K_s)\, ds | \mathcal{F}_t, \mathcal{C}_t\right] = V, \mathrm{E}[h|\mathcal{F}_t] = \phi \right\},$$

which represents the set of all contracts that are incentive compatible (as in Definition

1.2) and that provide the manager with continuation value V starting from beliefs ϕ.

To provide a definition of renegotiation-proof contract, I first define $\mathcal{J}(\mathcal{C})$ as the costs for the advisor that offers contract $\mathcal{C}$, and I define $\mathcal{O}(\mathcal{C}, t)$ as the time-t continuation contract implied by $\mathcal{C}$. Building on the definition in Strulovici (2011), I define a (weakly) renegotiation-proof contract as follows.

Definition 1.5 (Weakly Renegotiation-Proof Contract). *$\mathcal{C} \in \mathcal{I}(V_0, \phi_0, 0)$ is weakly renegotiation proof if, for all $t \geq 0$ and $t' \geq 0$ such that $V_t = V_{t'}$ and $\phi_t = \phi_{t'}$,*

$$\mathcal{J}(\mathcal{O}(\mathcal{C}, t)) = \mathcal{J}(\mathcal{O}(\mathcal{C}, t'))$$

In a renegotiation-proof contract, the advisor must be unable to renegotiate the terms of the contract in order to leave the manager indifferent and reduce costs for herself. In this environment, the advisor still fully commits to a promised value for the manager, but she is unable to commit to the manager's information rent. Therefore, the nature of the commitment in this model is similar to the "commitment and renegotiation" assumption in Laffont and Tirole (1990): Although the advisor and the manager can write a long-term contract, they can later agree to alter the terms of the initial contract if doing so is mutually beneficial.

Even when the advisor cannot commit to an information rent for the manager, enforcing full effort remains optimal. In other words, Proposition 1.2 remains valid because it does not rely on any particular assumption about the advisor's commitment. Moreover, (1.3) also remains valid because it characterizes incentive compatibility in any generic principal-agent setting.

For any given contract, the manager has an information rent given by equation

(1.11). However, unlike in the optimal contract with full commitment, the advisor does not take into account how the contract affects the agent's information rent. The advisor will instead take the process for the information rent as given and design a contract that is optimal given this process. This behavior reflects the fact that the manager and the advisor can mutually agree to change the terms of the contract in the future and modify the implied information rent.

We can refine the search for a renegotiation-proof contract by looking at the set of Markovian contracts, under the assumption that the optimal renegotiation-proof contract is unique given an initial promised value and a posterior.

Definition 1.6. *A contract $\mathcal{C} \in \mathcal{I}(V_0, \phi_0, 0)$ is Markovian if, for all $t \geq 0$ and $t' \geq 0$ such that $V_t = V_{t'}$ and $\phi_t = \phi_{t'}$, $\mathcal{O}(\mathcal{C}, t) = \mathcal{O}(\mathcal{C}, t')$*

Lemma 1.4. *Any Markovian contract is weakly renegotiation proof. If the optimal renegotiation-proof contract is unique, then it is Markovian.*

In this paper, I do not explore the issue of multiple optimal contracts. I simply focus on the optimal Markovian contract, which, given the previous lemma, is a renegotiation-proof contract. Consequently, given expression (1.11) for the manager's information rent and the Markovian structure of the contract, we also conclude that the manager's information rent is Markovian in the manager's continuation value and beliefs. Using the homogeneity of the problem, I can characterize an optimal renegotiation-proof contract as follows.

Proposition 1.8. *As before, let $\hat{v}_t = ((1 - \rho)V_t)^{\frac{1}{1-\rho}}$. In an optimal renegotiation-proof contract, we have that $C_t = \hat{v}_t c_R(\phi_t)$, $\beta_t = (1 - \rho)V_t \hat{\beta}_R(\phi_t)$ and $\eta_t = \eta_R(\phi_t)$,*

75

where $c_R(\phi), \hat{\beta}_R(\phi)$ and $\eta_R(\phi)$ are the optimal controls in the HJB equation:

$$rJ_R(\phi) = \min_{c,\hat{\beta},e} \left\{ \frac{c}{q(\eta)} - \mu(\eta,\phi,0)\frac{\hat{\beta} - \eta z_R(\phi)}{\sigma\lambda}c^\rho + J_R(\phi)\left(\frac{\delta}{1-\rho} - \frac{c^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}^2\right) \right.$$
$$\left. + \hat{\beta}\eta\phi(1-\phi)J_R'(\phi) + \frac{1}{2}\eta^2\phi^2(1-\phi)^2 J_R''(\phi) \right\}. \quad (1.27)$$

The function z_R solves the differential equation:

$$z_R(\phi)c_R(\phi)^{1-\rho} - \hat{\beta}_R(\phi)\eta_R(\phi)\phi(1-\phi) - (1-\rho)\hat{\beta}_R(\phi)\eta_R(\phi)\phi(1-\phi)z_R'(\phi)$$
$$= \frac{1}{2}\eta_R(\phi)^2\phi^2(1-\phi)^2 z_R''(\phi). \quad (1.28)$$

The cost function for the advisor is $\hat{v}_t J_R(\phi_t)$ and the manager's information rent is $\xi_t = (1-\rho)V_t z_R(\phi_t)$.

1.A.2 Market-Based Incentives

I now consider the case of a manager that directly collects money from investors. No explicit contracts are written; however, the manager can maintain a stake in the fund in order to ensure incentive compatibility. To maintain comparison with the problem that I have studied so far, I assume the manager's consumption is observable by the market, together with the amount of assets under management, the manager's stake, and his experimentation. Consequently, these quantities must be set as functions of the returns observed by the market, because the market could punish observed deviations by leaving the manager in autarky.

In this framework, the manager possesses some wealth, which he allocates between

a risk-free asset and the fund he runs. He raises capital from investors, who supply capital perfectly elastically at the risk-free rate r. Therefore, the manager can collect fees on the investor's capital for the difference between the expected return of his fund and the risk-free rate.

Absent an advisor who designs an explicit incentive contract, the manager must resort to other implicit incentives schemes to mitigate moral hazard problem. In particular, the manager holds a stake in the fund. If the manager had none of his wealth invested in the fund, he would be tempted to shirk to obtain private benefits. In this situation, rational investors would not be willing to provide capital. If, instead, the manager has some of his own wealth invested in the fund, then investors, to some extent, trust the manager with their own money. The stake of the manager should be interpreted in a broad sense. One way the manager can hold a stake in the fund is by being an investor in the fund. Alternatively, the manager may simply receive cash payments that depend on the performance of the fund and that are not immediately consumed. Such an example is a manager who charges symmetric (fulcrum) fees for performance.

As in the principal-agent formulation, Proposition (1.2) continues to hold and the manager will choose a strategy that will credibly enforce full effort. The proof of Proposition (1.2) holds by simply reinterpreting the advisor's cost function as the wealth of the manager. However, the incentive-compatibility condition needs to be slightly reformulated to fit the new contractual environment.

The incentive-compatibility condition remains essentially unchanged: Changes in the continuation value of the manager, keeping the manager's beliefs fixed, should

exceed the marginal utility of shirking, as in equation (1.8). However, because no advisor is committing to a promised value for the manager, the implementation of this incentive-compatibility condition has to rely on the dynamics of the manager's wealth and beliefs.

In equilibrium, the manager's continuation value is given by a function $V(A, \phi)$, which depends on the manager's wealth A and the equilibrium beliefs about his skill ϕ. The manager's wealth evolves as

$$dA_t = \left(rA_t - \frac{C_t}{q(\eta_t)} + \mu(\eta_t, \phi_t, 0)K_t \right) dt + \Theta_t \sigma dW_t^{\mathcal{C}},$$

where Θ is the manager's stake. The manager collects fees on the capital he raised from investors. Fees account for a part $\mu(\phi_t, \eta_t, 0)(K_t - \Theta_t)$ of the manager's instantaneous cash flow. The remaining part, $\mu(\phi_t, \eta_t, 0)\Theta_t$, is the expected excess return from his stake in the fund. In an incentive-compatible contract, fees and expected returns coincide and the stake Θ_t does not affect expected cash flows, but only the volatility of cash flows.

I therefore rewrite the incentive-compatibility condition (1.8) by using the Markovian characterization of the manager's continuation value.

Lemma 1.5. *If the market-based allocation is incentive compatible,*

$$V_A(A_t, \phi_t)\Theta_t \sigma + V_\phi(A_t, \phi_t)\eta_t \phi_t(1 - \phi_t) \geq u'(C_t)K_t \lambda \sigma + \eta_t \xi_t, \tag{1.29}$$

where ξ_t evolves as

$$d\xi_t = [\delta\xi_t - \eta_t\phi_t(1 - \phi_t)(V_A(A_t, \phi_t)\Theta_t\nu(\eta_t) + V_\phi(A_t, \phi_t)\eta_t\phi_t(1 - \phi_t))]dt + \delta_t dW_t.$$

This incentive-compatibility condition includes a role for the career concerns of the manager, which were initially studied by Fama (1980) and Holmström (1999). The incentive-compatibility condition (1.29) captures the result in Gibbons and Murphy (1992) that direct incentives and career concerns add up in determining the total incentives of the manager. In my model, direct incentives are represented by the value of the manager's stake, $V_A(A_t, \phi_t)\Theta_t\sigma$. Career concerns are represented by the difference $V_\phi(A_t, \phi_t)\eta_t\phi_t(1 - \phi_t) - \eta_t\xi_t$. To see why, recall that $V_\phi(A_t, \phi_t)$ is the marginal value of equilibrium beliefs, which are common among the market and the manager. From the discussion in section 1.3.4, the information rent ξ_t is the product between $\phi_t(1 - \phi_t)$ and the marginal value of the manager's private beliefs. Therefore, the difference

$$V_\phi(A_t, \phi_t) - \frac{\xi_t}{\phi_t(1 - \phi_t)}$$

represents the marginal value of the market's beliefs. The manager prefers the market's beliefs to be high in order to collect higher fees. The incentive-compatibility condition (1.29) can then be written as

$$\underbrace{V_A(A_t, \phi_t)\Theta_t\sigma}_{\text{direct incentives}} + \underbrace{\eta_t\phi_t(1 - \phi_t)\left(V_\phi(A_t, \phi_t) - \frac{\xi_t}{\phi_t(1 - \phi_t)}\right)}_{\text{career concerns}} \geq u'(C_t)K_t\lambda\sigma,$$

which highlights that both direct incentives and career concerns play a role in pro-

79

viding incentives to the manager.

Because the manager and the investors interact in a spot market, the implicit market-based contract must also be renegotiation proof. The manager's information rent will therefore be a function of his wealth and equilibrium beliefs,

$$\xi_t = \xi(A_t, \phi_t).$$

Given the functional form of the utility function, the manager's continuation value and information rent are homogeneous in wealth, that is, $V(A, \phi) = \frac{A^{1-\rho}}{1-\rho} v_M(\phi)$ and $\xi(A, \phi) = \frac{A^{1-\rho}}{1-\rho} z_M(\phi)$. Therefore, the incentive-compatibility constraint can be written as

$$(1 - \rho)v_M(\phi_t)\hat{\Theta}_t\sigma + v'_M(\phi_t)\eta_t\phi_t(1 - \phi_t) \geq (1 - \rho)c_t^{-\rho}k_t\lambda\sigma + \eta_t z_M(\phi_t),$$

where I use $\hat{\Theta}_t = \frac{\Theta_t}{A_t}$, $c_t = \frac{C_t}{A_t}$ and $k_t = \frac{K_t}{A_t}$.

This contracting environment can thus be characterized as follows.

Proposition 1.9. *In an optimal market-based contract, $C_t = A_t c_M(\phi_t)$, $\Theta_t = A_t \hat{\Theta}_M(\phi_t)$ and $\eta_t = \eta_M(\phi_t)$, where $c_M(\phi), \hat{\Theta}_M(\phi)$ and $\eta_M(\phi)$ are the optimal con-*

80

trols in the HJB equation

$$\delta v_M(\phi) = \max_{c,\Theta,\eta,k} \left\{ c^{1-\rho} + v_M(\phi)(1-\rho) \left[r - \frac{c}{q(\eta)} + \mu(\eta,\phi,0)k - \frac{1}{2}\rho\hat{\Theta}^2\sigma^2 \right] \right.$$

$$\left. + (1-\rho)\sigma\eta\hat{\Theta}\phi(1-\phi)v_M'(\phi) + \frac{1}{2}\eta^2\phi^2(1-\phi)^2 v_M''(\phi) \right\}$$

(1.30)

$$\text{s.t. } (1-\rho)v_M(\phi)\hat{\Theta}\sigma + v_M'(\phi)\eta_t\phi(1-\phi) = (1-\rho)c^{-\rho}k\lambda\sigma + \eta z_M(\phi). \quad (1.31)$$

The function z_M solves the differential equation

$$\left[\delta - (1-\rho)\left(r - \frac{c_M(\phi)}{q(\eta_M(\phi))} + \mu(\eta_M(\phi),\phi,0)k_M(\phi) - \frac{1}{2}\rho\hat{\Theta}(\phi)^2\sigma^2 \right) \right] z_M(\phi) =$$

$$(1-\rho)\eta_M(\phi)\phi(1-\phi)\hat{\Theta}(\phi)\sigma v_M(\phi) + \eta_M(\phi)^2\phi^2(1-\phi)^2 v_M'(\phi)$$

$$+ (1-\rho)\eta_M(\phi)\phi(1-\phi)\hat{\Theta}(\phi)\sigma z_M'(\phi) + \frac{1}{2}\eta_M(\phi)^2\phi^2(1-\phi)^2 z_M''(\phi). \quad (1.32)$$

The value function for the manager is $\frac{A_t^{1-\rho}}{1-\rho}v(\phi_t)$ *and his information rent is*
$\xi_t = \frac{A_t^{1-\rho}}{1-\rho}z_M(\phi_t)$.

1.A.3   Results

I introduced the two models of this section separately in order to highlight the difference between the two contracting environments. However, the two models are, as the next proposition shows, outcome equivalent. As long as the manager has the same commitment power of the advisor and as long as the market can observe the same actions that the advisor can observe, a model with market-based incentives is
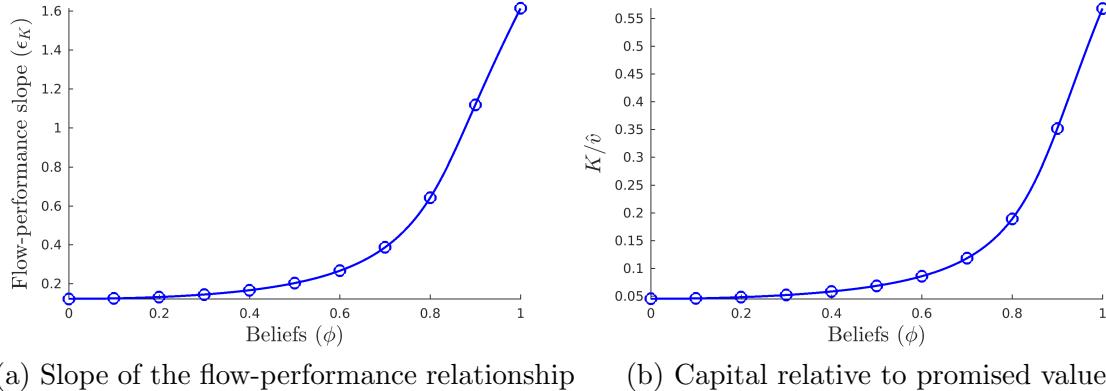
(a) Slope of the flow-performance relationship   (b) Capital relative to promised value

Figure 1.8: Slope of the flow-performance relationship and capital in the optimal renegotiation-proof contract. The parameters of the model are as in Table 1.1.

equivalent to a principal-agent model. In particular, the two models imply the same social welfare and the same outcomes in terms of flows and compensation.

Proposition 1.10. *The renegotiation-proof contract and the market-based contract are equivalent.*

In Appendix 1.E, I provide a verification argument based on the recursive representations of the models in Propositions 1.8 and 1.9. However, we can immediately observe that the renegotiation-proof contracting problem coincides with the problem of a manager who wants to find the minimal wealth that could support a given lifetime utility with market-based incentives.

We can now study the flow-performance relationship and the compensation scheme in these contractual environments. As the figures in this section show, the results are qualitatively identical to the full-commitment model with only one exception.[14]

---

14. I numerically solve for the models using the same parameters of section 1.5.
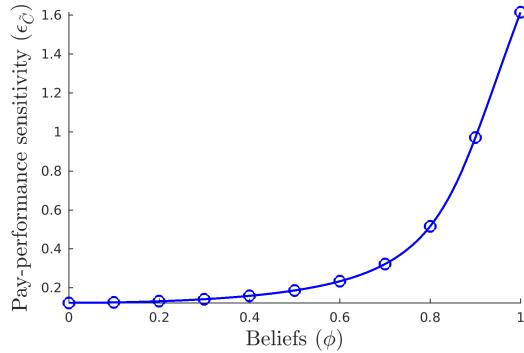
(a) Pay-performance sensitivity
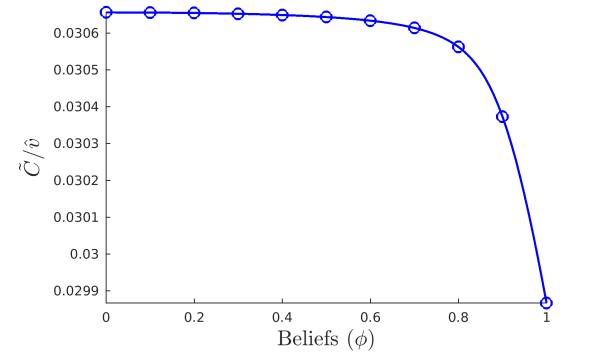
(b) Compensation relative to promised value

Figure 1.9: Pay-performance sensitivity and compensation in the optimal renegotiation-proof contract. The parameters of the model are as in Table 1.1.

Because the advisor or the manager is unable to commit to ex post inefficient actions, the flow-performance relationship and the pay-performance sensitivity depend only on the current beliefs ϕ. In the full-commitment model, the advisor could commit to reduce future information rents through a multiplier y with a negative drift. As a result, over time, the advisor would change the manager's incentives and experimentation, even for the same value of beliefs ϕ. This channel of non-stationarity is absent when the advisor cannot commit to an information rent for the manager.

As before, the flow-performance relationship is positive and increasing in beliefs ϕ (Figure 1.8a). A good return is associated not only with a positive flow of capital into the fund, but also with an increase in the slope of the flow-performance relationship, because beliefs ϕ increase after good performance. The economic motivation is identical to the one I discussed in section 1.5.2 and is related to the willingness of the advisor (or the manager) to increase capital, relative to the manager's risk tolerance, when beliefs are higher (Figure 1.8b).

83

(a) Explicit incentives $\hat{\beta}_R$ (b) Market-based incentives $\hat{\Theta}_M$
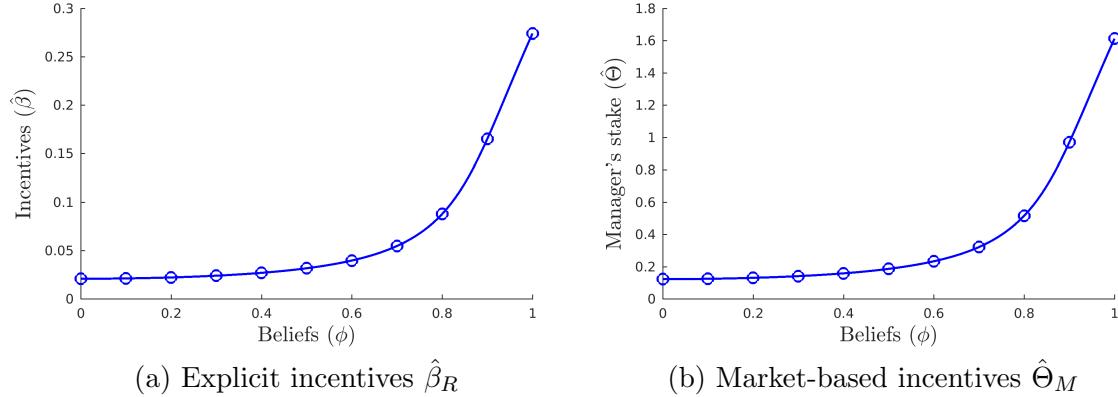
Figure 1.10: Incentives in the optimal renegotiation-proof contract and in the market-based contract. The parameters of the models are as in Table 1.1.

The relation between compensation and performance also matches the full-commitment contract. Compensation increases with returns, and the pay-performance sensitivity increases with past performance (Figure 1.9a), which then results in a convex relation between cumulative pay and cumulative performance. Moreover, performance-based compensation is again back-loaded after positive returns (Figure 1.9b). Interestingly, the manager does not need an advisor to defer his compensation and increase his stake in the fund. This result is consistent with anecdotal evidence that hedge fund managers tend to invest most of their performance fees in the hedge fund itself (Agarwal et al., 2009).

Although the outcomes of the two models are identical, the implementation is different. In the contract designed by the advisor, the advisor adjusts the explicit contractual terms of the manager to ensure incentive compatibility (Figure 1.10a). In the market-based allocation, the manager adjusts his implicit incentives by holding a larger stake in the fund (Figure 1.10b). In the optimal mechanisms, these two im-

plementations will provide the manager with identical incentives and, consequently, result in an identical flow-performance relationship and compensation scheme.

Appendix 1.B   Static Model

In this appendix, I consider a two-period contracting model in order to provide an analytically tractable example of how, when moral hazard interacts with uncertainty, a principal optimally designs a contract that features reduced experimentation in the future.

To improve tractability, the assumptions on the agent's preferences are different from the main model. However, rather than being a shortcoming of this example, this difference further goes to the generality of my result.

Before studying the two-period model, consider a one-period model to establish a benchmark. Suppose returns are normally distributed and given by

$$R \sim N(\alpha + \eta h, \sigma^2) - m,$$

where m is the agent's hidden action, which gives him a private benefit λm, and $h \in \{0, 1\}$ is the agent's unknown skill. The principal is risk neutral, whereas the agent has CARA utility with absolute risk aversion γ. Because returns are normal, the agent's utility can be expressed in the mean-variance form. I focus on linear contracts that enforce $m = 0$. In these contracts, the agent's final consumption is given by

$$V = \Delta + C,$$

where

$$\Delta = \beta(R - (\alpha + \eta\phi_0))$$

captures the risk exposure of the agent, and C represents his expected consumption.

The principal controls experimentation through η. I assume that $\eta \in \{0, 1\}$ and that it involves no costs. To ensure incentive compatibility, we must impose $\beta \geq \lambda$. Then, the optimal contract solves

$$\max \mathrm{E}[R] - C$$

$$\text{s.t. } \beta \geq \lambda$$

$$C - \frac{\gamma}{2}\beta^2 \geq U_0.$$

It is straightforward to verify the following lemma:

Lemma 1.6. *In a one-period contract, $\beta = \lambda$ and $\eta_1 = \mathbb{I}\{\phi_0 \geq 0\}$.*

Now consider a two-period model. Output in period t is normally distributed and given by

$$R_t \sim N(\alpha + \eta_t h, \sigma^2) - m_t,$$

where m_t is the hidden action of the agent, which gives him a private benefit λa_t, with $\lambda \in (0, 1)$. To reduce unnecessary complications, I assume $\eta_0 = 1$ and $\phi_0 > 0$. The principal can choose $\eta_1 \in \{0, 1\}$. Because $m_t > 0$ is inefficient, I focus on contracts that implement $m_t = 0$.

Neither the principal or the agent discount the future. The principal is risk neutral, whereas the agent has CARA preferences over the final payout. The final

86

compensation of the agent is given by

$$V = \Delta_0 + \Delta_1 + C,$$

where Δ_t is the performance-based compensation at time t, and C is a promised expected compensation that is pinned down by a participation constraint.

I focus on linear contracts where Δ_t takes the form $\Delta_t = \beta_t(R_t - (\alpha + \eta_t \phi_t))$, and where β_t is chosen in order to make the contract incentive compatible. Because the agent is risk averse, in an optimal contract, β_t will be chosen to be as small as possible, as long as it ensures incentive compatibility.

Finally, once the principal and the agent sign a two-period contract, they fully commit to it. In particular, the principal can commit to a future choice of η_1.

It is immediate to notice that $\beta_1 = \lambda$, because the last stage of the game is analogous to a standard static problem. However, the choice of β_0 depends on future learning. In particular, to ensure incentive compatibility, me must have

$$\beta_0 \geq \lambda\left(1 - \mathrm{E}[\eta_1 \phi_m(R_0, 0)]\right),$$

where $\phi(R_0, m)$ is the principal's posterior as a function of the time 0 return R_0 and the time 0 hidden action m.

By Bayes' Law, the posterior is given by

$$\phi_1 = (1 - \bar{v})\phi_0 + \bar{v}(R_0 - \alpha),$$

so that the constraint on β_0 becomes

$$\beta_0 \geq \lambda + \lambda \bar{v} \mathrm{E}[\eta_1].$$

This inequality is the counterpart of (1.8) in the paper, and $\lambda \bar{v} \mathrm{E}[\eta_1]$ represents the agent's information rent in this two-period model.

Similar to the one-period model, the principal solves

$$\max \mathrm{E}[R_0 + R_1] - C$$

$$\text{s.t. } \beta_0 \geq \lambda + \lambda \bar{v} \mathrm{E}[\eta_1]$$

$$C - \frac{\gamma}{2}(\beta_0^2 + \lambda^2) \geq U_0.$$

Because the incentive-compatibility and participation constraints must bind, the problem reduces to finding an R_0-measurable learning strategy η_1 that maximizes

$$\mathrm{E}[\eta_1(\phi_1 - \gamma\lambda^2\bar{v})] - \frac{\gamma}{2}\lambda^2\bar{v}^2(\mathrm{E}[\eta_1])^2.$$

Unlike the one-period model, setting $\eta_1 = 1$ when expected returns are positive is no longer always optimal.

Proposition 1.11. *There exists a $\bar{\phi}$ such that $\eta_1 = 0$ if $\phi_1 < \bar{\phi}$, whereas $\eta_1 = 1$ if $\phi_1 > \bar{\phi}$. Moreover, $\bar{\phi} > 0$.*

Proof. Let a state be a return realization R at time 0. Suppose that there exist a set of states RR and a set of states RR' such that the posterior ϕ_1 in all states in RR is lower than the posterior in all states of RR', and such that also that $\eta_1 = 1$ in

88

RR, while $\eta_1 = 0$ in RR'. Let ν be the measure associated with the normal density with mean $\alpha + \phi_0$ and variance σ^2. Consider $n = min(\nu(RR), \nu(RR'))$ and suppose that $n = \nu(RR')$. Now consider a set $rr \subset RR$ and a set $rr' \subset RR'$ such that $\nu(rr) = \nu(rr') = n$. Consider a new contract with a new experimentation policy η'. This policy is such that: (i) if $R \in rr'$, $\eta'(rr') = 1$, (ii) if $R \in rr$, $\eta(R) = 0$ (iii) if $R \in RR' \setminus rr'$, $\eta'(R) = 0$, and (iii) if $R \in RR \setminus rr$, $\eta'(R) = 1$. Since rr' and rr have the same measure, it follows that $(E[\eta_1'])^2 = (E[\eta_1])^2$. However, since $\phi_1(R') \geq \phi_1(R)$ for all $R' \in rr'$ and $R \in rr$, under the new policy we have $E[\eta_1'(\phi_1 - \gamma\lambda^2\bar{v})] > E[\eta_1(\phi_1 - \gamma\lambda^2\bar{v})]$. This contradicts η_1 being optimal. Given that η_1 is binary, it immediately follows that the optimal choice takes the form of a threshold rule. Denote the threshold with $\bar{\phi}$.

It remains to show that the threshold $\bar{\phi}$ is positive. To do so, consider the objective function, which can be written as

$$E\left[\eta_1\left(\phi_1 - \gamma\lambda^2\bar{v} - \frac{\gamma}{2}\lambda^2\bar{v}^2 E[\eta_1]\right)\right].$$

The necessary condition for optimality is that

$$\eta_1 = 0 \quad \text{if} \quad \phi_1 \leq \gamma\lambda^2\bar{v} + \frac{\gamma}{2}\lambda^2\bar{v}^2 E[\eta_1],$$

which then implies that $\bar{\phi} \geq \gamma\lambda^2\bar{v} + \frac{\gamma}{2}\lambda^2\bar{v}^2 E[\eta_1] > 0$.

$\square$

As in the dynamic model of the paper, the principal commits to a reduced amount of future learning in order to improve current incentives. The mechanism in this static

model is very clear: Future experimentation η_1 requires stronger incentives β_0. But because exposing the agent to risk is costly, the principal has an incentive to reduce future experimentation below the ex post optimal level in order to improve ex ante incentives.

Appendix 1.C   Proofs for Section 1.3

Before proceeding to the proofs, I need to introduce some formal notation that I have omitted in the main body of the paper. Let $(\Omega, \mathcal{F}^*, P)$ be a probability space. $(W_t)_{t \geq 0}$ is a Wiener process on this probability state and the manager's skill h is a random variable on $(\Omega, \mathcal{F}^*)$. The path of returns $(R_t)_{t \geq 0}$ is also a random variable $(\Omega, \mathcal{F}^*)$ on $(\Omega, \mathcal{F}^*)$. I denote with $P^{\mathcal{C}}$ a probability measure over the set of paths of returns for a given contract $\mathcal{C}$. As in section 1.3.3, $(\mathcal{F}_t)_{t \geq 0}$ is the filtration generated by the path of returns $(R_t)_{t \geq 0}$, possibly augmented by the P-null sets.

1.C.1   Proof of Proposition 1.1

Proof. Let P^G be the probability measure on $(\Omega, \mathcal{F}_\infty)$ conditional on $h = 1$ and let P^B be the probability measure on $(\Omega, \mathcal{F}_\infty)$ conditional on $h = 0$.

Let

$$W_t^{\mathcal{C},0} = \int_0^t dR_s - (r + \mu(\eta_s, 0, m_s))ds \tag{1.33}$$

and define the likelihood ratio

$$X_t = \exp\left\{\int_0^t \eta_s \, dW_s^{\mathcal{C},0} - \frac{1}{2}\int_0^t \eta_s^2 \, ds\right\},$$

which represents the ratio between the likelihood that the path $(R_s)_{0 \leq s \leq t}$ is generated by a skilled manager $(h = 1)$ and the likelihood that the same path is generated by an unskilled manager $(h = 0)$.

Since players use Bayes' rule to form beliefs,

$$\phi_t = \frac{pX_t}{pX_t + (1-p)}. \tag{1.34}$$

We can then apply Ito's lemma and obtain

$$d\phi_t = -\frac{(1-p)p^2}{(pX_t + (1-p))^3}(\eta_t X_t)^2 dt + \frac{(1-p)p}{(pX_t + (1-p))^2}\eta_t X_t dW_t^{\mathcal{C},0}.$$

Using (1.33) and (1.34), we conclude that

$$d\phi_t = \eta_t(1 - \phi_t)\phi_t \frac{1}{\sigma}(dR_t - (r + \mu(\eta_t, \phi_t, m_t))dt).$$

$\square$

1.C.2   Proof of Proposition 1.2

I proceed by contradiction. Let $\mathcal{C}$ be an optimal contract and let $(m_t)_{t \geq 0}$ be the best response of the manager to the contract. Suppose that $(m_t)_{t \geq 0}$ is strictly larger than zero with positive probability.

Since this contract is $\mathcal{F}_t$-adapted, this means that the contract specifies the time t allocation (C_t, K_t, η_t, m_t) as a function of the history of returns $(R_s)_{0 \leq s \leq t}$. For

example, compensation at time t can be written as,

$$C_t = \mathbf{C}^t((R_s)_{0 \leq s \leq t}),$$

experimentation at time t can be written as

$$\eta_t = \boldsymbol{\eta}^t((R_s)_{0 \leq s \leq t}).$$

and shirking at time t can be written as

$$m_t = \mathbf{m}^t((R_s)_{0 \leq s \leq t}).$$

Now consider an alternative contract $\hat{\mathcal{C}}$. This contract is designed in the following way. If the history of returns at time t is $(R_s)_{0 \leq s \leq t}$, the contract specifies capital and experimentation at time t as equal to the capital and experimentation that contract $\mathcal{C}$ specifies after history $(R_s - \int_0^s \hat{m}_u \, du)_{0 \leq s \leq t}$, where

$$\hat{m}_t = \mathbf{m}^t\left(\left(R_s - \int_0^s \hat{m}_u \, du\right)_{0 \leq s \leq t}\right).$$

For example, experimentation at time t for contract $\hat{C}$ is given by

$$\hat{\eta}_t = \boldsymbol{\eta}^t\left(\left(R_s - \int_0^s \hat{m}_u \, du\right)_{0 \leq s \leq t}\right).$$

However, under contract $\hat{\mathcal{C}}$, consumption is specified as

$$\hat{C}_t = \mathbf{C}^t \left(\left(R_s - \int_0^s \hat{m}_u \, du \right)_{0 \le s \le t} \right) + \lambda \hat{m}_t \hat{K}_t.$$

If the agent never shirks when he's offered the alternative contract $\hat{\mathcal{C}}$, he obtains a consumption process that coincides with the consumption process he obtains by shirking in contract $\mathcal{C}$. If the agent chooses a shirking process $(m'_t)_{t \ge 0}$ with contract $\hat{\mathcal{C}}$, he obtains the same consumption process he would have obtained in contract $\mathcal{C}$ from a shirking process $(m_t + m'_t)_{t \ge 0}$.

Because $(m_t)_{t \ge 0}$ is the best response of the manager to contract $\mathcal{C}$, that is,

$$(m_t)_{t \ge 0} \in \arg\max_{(m'_t)_{t \ge 0}} \, \mathrm{E}\left[\int_0^\infty e^{-\delta t} u(C_t + m'_t \lambda K_t) \, dt \,\Big|\, \mathcal{F}_0, \mathcal{C} \right]$$

then it must be the case that $(0)_{t \ge 0}$ is the best response of the manager to contract $\hat{\mathcal{C}}$, that is,

$$(0)_{t \ge 0} \in \arg\max_{(m'_t)_{t \ge 0}} \, \mathrm{E}\left[\int_0^\infty e^{-\delta t} u(\hat{C}_t) \, dt \,\Big|\, \mathcal{F}_0, \hat{\mathcal{C}} \right],$$

since, otherwise, $(m_t)_{t \ge 0}$ would not be a best response in the original contract.

Under the new contract $\hat{\mathcal{C}}$ the agent receives the same lifetime utility as in contract $\mathcal{C}$. However, the costs for the principal change by

$$\mathrm{E}\left[\int_0^\infty e^{-rt} \left(1 - \frac{\lambda}{q(\eta_t)} \right) m_t K_t \, dt \right] \le 0.$$

Since $q(\bar{\eta}) \ge \lambda$, then the principal is now bearing lower costs. This contradicts the assumption that $\mathcal{C}$ is an optimal contract.

1.C.3 Proof of Proposition 1.3

I use the stochastic maximum principle to derive necessary conditions for incentive compatibility, as in (Williams, 2011). Fix an incentive compatible contract with full effort, $\mathcal{C}$. Consider a shirking process $m = (m_t)_{t \geq 0}$. Let $P^{\mathcal{C}}$ be the probability measure over path of returns for which $W_t^{\mathcal{C}}$ is a standard Brownian motion. Let P^m be a measure under which

$$W_t^m = W_t^{\mathcal{C}} - \int_0^t \frac{m_t}{\sigma} \, dt$$

is a standard Brownian motion.

I denote with τ a stopping time at which the principal stops experimentation forever, that is, $\tau = \inf\{t \geq 0 : \eta_s \leq 0, \forall s \geq t\}$. After time τ, Lemma 1.3 applies and the skill of the manager becomes irrelevant for returns and incentive provision. In particular, the manager's continuation value at τ, V_τ, does not depend on beliefs about the manager's skill.

Instead of considering the posterior beliefs ϕ_t as state variable, it is convenient to work with the log-likelihood ratio $x_t = \log X_t$, where X_t is the likelihood ratio. We can then express beliefs as a function of the log-likelihood ratio,

$$\psi(x) = \frac{pe^x}{1 - p + pe^x}. \tag{1.35}$$

More precisely, let x_t be the principal's log-likelihood ratio and let $x_t + \Delta x_t$ be

the agent's log-likelihood ratio. The laws of motion of x_t and Δx_t are given by:

$$dx_t = \left(\psi(x_t) - \frac{1}{2}\right)\eta_t^2 dt + \eta_t dW_t^{\mathcal{C}}$$

$$d\Delta x_t = \frac{m_t}{\sigma}\eta dt.$$

The continuation value of the agent, given $\mathcal{C}$ and m, can be written as

$$\mathrm{E}\left[\int_0^\infty \Gamma_t^m e^{-\delta t} u(C_t + m_t \lambda K_t)\, dt \Big| \mathcal{F}_0\right].$$

where $\Gamma^m \equiv \frac{dP^a}{dP^{\mathcal{C}}}$ is a density process representing the change of measure for the path of returns induced by the shirking strategy m. By Girsanov's Theorem, Γ^m as

$$d\Gamma_t^m = \left(-\frac{m_t}{\sigma} + \left(\psi(x_t + \Delta x_t) - \psi(x_t)\right)\right)\Gamma_t dW_t^{\mathcal{C}}.$$

Let $\mathcal{V}$ be the multiplier for Γ^m and let β_t be the multiplier's volatility. Similarly, let ξ_t be the multiplier for Δx_t with volatility ω_t. The Hamiltonian for the agent's optimization problem is the following:

$$\Gamma u(C + m\lambda K) + \left(-\frac{m}{\sigma} + \left(\psi(x + \Delta x) - \psi(x)\right)\right)\Gamma\beta + \frac{m}{\sigma}\eta\xi \qquad (1.36)$$

If $m = 0$ is optimal, it must be the case that $m = 0$ maximizes (1.36) with $\Gamma^m = 1$

and $\Delta x = 0$, which happens only if

$$u'(C)\lambda K + \frac{\eta\xi}{\sigma} - \frac{\beta}{\sigma} \leq 0.$$

The multipliers $\mathcal{V}$ and ξ solve the following backward stochastic differential equations (BSDEs)

$$d\mathcal{V}_t = (\delta\mathcal{V}_t - u(C_t))dt + \beta_t dW_t^{\mathcal{C}}$$

$$d\xi_t = (\delta\xi_t - \phi_t(1 - \phi_t)\eta_t\beta_t)dt + \delta_t dW_t^{\mathcal{C}}, \tag{1.37}$$

with terminal conditions

$$\mathcal{V}_\tau = V_\tau$$

$$\xi_\tau = 0.$$

Solving the BDSE for $\mathcal{V}_t$, we obtain

$$\mathcal{V}_t = V_t = \mathrm{E}\left[\int_t^\infty e^{-\delta t}u(C_t)\, dt \middle| \mathcal{F}_t\right].$$

1.C.4   Proof of Proposition 1.4

Proof. To show that $\xi_t = \phi_t(1 - \phi_t)\partial_\phi V_t$, it is sufficient to show that $\xi_t = \partial_x V_t$, where x_t is the log-likelihood ratio at time t. The fact that $\partial_x V_t = \phi_t(1 - \phi_t)\partial_\phi V_t$ follows from equation (1.35).

Consider

$$V_t = \mathrm{E}\left[\int_t^\infty e^{-\delta s}u(C_s)\, ds \middle| \mathcal{F}_t\right].$$

96

Given an initial log-likelihood ratio x_t, we can write

$$(1 - p + pe^{x_t})V_t = \mathrm{E}\left[\int_t^\infty (1 - p + pe^{\Delta x_s + x_t})e^{-\delta s}u(C_s)\,ds\Big|\mathcal{F}_t, h = 0\right].$$

Differentiating with respect to x_t we obtain

$$pe^{x_t}V_t + (1 - p + pe^{x_t})\partial_x V_t = \mathrm{E}\left[\int_t^\infty pe^{\Delta x_s + x_t}e^{-\delta s}u(C_s)\,ds\Big|\mathcal{F}_t, h = 0\right], \quad (1.38)$$

which can be rearranged as

$$\partial_x V_t = \phi_t(G_t - V_t)$$

where

$$G_t = \mathrm{E}\left[\int_t^\infty e^{-\delta s}u(c_s)\,ds\Big|\mathcal{F}_t, h = 1\right].$$

Using Girsanov's theorem and Ito's lemma, we can derive the law of motion of $\phi_t(G_t - V_t)$,

$$d\phi_t(G_t - V_t) = (\delta\phi_t(G_t - V_t) - \eta_t\beta_t\phi_t(1 - \phi_t))dt + \omega_t'dW_t^{\mathcal{C}}, \quad (1.39)$$

for some $\mathcal{F}_t$-adapted process $(\omega_t')_{t\geq 0}$, with the terminal condition $\phi_\tau(G_\tau - V_\tau) = 0$, where $\tau = \inf\{t \geq 0 : \eta_s \leq 0, \forall s \geq t\}$. Using the comparison principal for BSDEs (1.37) and (1.39) (Pham, 2009), we conclude that $\xi_t = \phi_t(G_t - V_t) = \partial_x V_t$. Moreover, we can solve the BSDEs and obtain

$$\xi_t = \mathrm{E}\left[\int_t^\infty e^{-\delta s}\eta_s\beta_s\phi_s(1 - \phi_s)\,ds\Big|\mathcal{F}_t\right].$$

It now remains to prove that, in an incentive compatible contract, $\xi_t \geq 0$. In order to show that ξ_t is positive, consider the BSDE:

$$\hat{\xi}_t = (\delta \hat{\xi}_t - \eta_t^2 \phi_t (1 - \phi_t)\xi_t - \eta_t \phi_t (1 - \phi_t)u'(C_t)\lambda K_t \sigma)dt + \hat{\sigma}dW_t^{\mathcal{C}},$$

with terminal condition $\hat{\xi}_\tau = 0$.

By the incentive-compatibility condition (1.8), $-\eta_t \xi_t - u_c(C_t)K_t \lambda \sigma \geq -\beta_t$. Moreover, $\eta_t \phi_t (1 - \phi_t) \geq 0$. Hence, from the comparison principle for BSDE (Pham, 2009), it follows that $\hat{\xi}_t \leq \xi_t$. Moreover, $\hat{\xi}_t$ can be written in closed form as

$$\hat{\xi}_t = \mathrm{E}\left[\int_t^\infty e^{-\int_0^s \delta - \eta_u^2 \phi_u(1-\phi_u)\,du} \eta_s \phi_s (1 - \phi_s)u'(C_s)\lambda K_s \sigma\,ds \Big| \mathcal{F}_t\right] \geq 0,$$

from which we conclude that $\xi_t \geq 0$. $\qquad\square$

1.C.5 Proof of Proposition 1.5

Proof. Let $\tilde{\phi}_t$ be the agent's posterior at time t, while ϕ_t is the principal's. Define

$$\zeta_t = \frac{\xi_t}{\phi_1(1 - \phi_t)},$$

which evolves as

$$d\zeta_t = \left(\delta \zeta_t - \eta_t \beta_t + \eta_t^2 \phi_t(1 - \phi_t)\zeta_t - \eta_t(1 - 2\phi_t)\omega_{\zeta t}\right)dt + \omega_{\zeta t}dW_t^{\mathcal{C}}$$

where

$$\omega_{\zeta t} = \frac{\omega_t - \xi_t(1 - 2\phi_t)\eta_t}{\phi_t(1 - \phi_t)}. \tag{1.40}$$

I want to show that, if the conditions of the proposition are satisfied, then $V_t + (\tilde{\phi}_t - \phi_t)\zeta_t$ is an upper bound on the continuation value of the agent at time t. Since $\tilde{\phi}_0 = \phi_0$, this will prove that the agent has no (strictly) better strategy than choosing $m_t = 0$ for all $t \geq 0$.

Let $\tau = \inf\{t \geq 0 : \eta_s \leq 0, \forall s \geq t\}$ be the stopping time at which the principal stops experimentation. For $t \geq \tau$ we have that $\zeta_t = 0$ and standard arguments imply that V_t is an upper bound for the agent's continuation value (DeMarzo and Sannikov, 2006; Sannikov, 2008).

For $t < \tau$, consider an arbitrary deviation up to time t and let

$$G_t = \int_0^t e^{-\delta s} u(C_s + m_s \lambda K_s)\, ds + e^{-\delta t}\left(V_t + (\tilde{\phi}_t - \phi_t)\zeta_t\right).$$

It suffices to show that G_t is a supermartingale for $t < \tau$. Indeed, in this case

$$G_t \geq \mathrm{E}[G_\tau | \mathcal{F}_t] = \mathrm{E}\left[\int_0^\tau e^{-\delta s} u(C_s + m_s \lambda K_s)\, ds + e^{-\delta \tau} V_\tau \Big| \mathcal{F}_t\right],$$

which then would imply that

$$V_t + (\tilde{\phi}_t - \phi_t)\zeta_t \geq \mathrm{E}\left[\int_t^\tau e^{-\delta(s-t)} u(C_s + a_s \lambda K_s)\, ds + e^{-\delta \tau} V_\tau \Big| \mathcal{F}_t\right].$$

In order to show that W_t is a supermartingale, it is sufficient to prove that the drift of dW_t is non-positive. Using Ito's lemma, and after some simplifications, the

99

drift of dG_t reduces to

$$e^{-\delta t}\left[u(C_t + m_t\lambda K_t) - u(C_t) - \beta_t\frac{m_t}{\sigma} + \phi_t(1-\phi_t)\zeta_t\frac{m_t}{\sigma} - (\tilde{\phi}_t - \phi_t)\omega_{\zeta t}\frac{m}{\sigma}\right].$$

Recall that $\xi_t = \phi_t(1-\phi_t)\zeta_t$. Combining the assumptions of the proposition with equation (1.40), we conclude that the drift is maximized for $m_t = 0$. For $m_t = 0$ the drift of dG_t is zero. Hence, for an arbitrary deviation m_t, drift of G_t is non-positive and G_t is a supermartingale. $\qquad\square$

Appendix 1.D   Proofs for Section 1.4

To prove the two propositions in Section 1.4, I first develop a series of lemmas to help organize the results. Before presenting the lemmas, let me introduce some notation that will facilitate the exposition. I denote with $\mathcal{C}^P_{V,\xi,\phi}$ the optimal contract for the primal problem when the initial state is given by V, ξ and ϕ. I denote with $\mathcal{C}^D_{V,Y,\phi}$ the optimal contract for the dual problem when the initial state is given by V, Y and ϕ.

Let $\mathcal{C}$ be an incentive-compatible contract that enforces no shirking, delivers expected lifetime utility V to the agent and that implies an information rent ξ. I denote with $J(V,\xi,\phi|\mathcal{C})$ the primal cost function when the principal chooses contract $\mathcal{C}$. Similarly, I denote with $G(V,Y,\phi|\mathcal{C})$ be the dual cost function when the principal chooses contract $\mathcal{C}$. Then we must have that $J^*(V,\xi,\phi) = J(V,\xi,\phi|\mathcal{C}^P_{V,\xi,\phi})$ and $G^*(V,Y,\phi) = G(V,Y,\phi|\mathcal{C}^D_{V,Y,\phi})$.

Any optimal contract for the dual problem implies an information rent. Given

a contract $\mathcal{C}$ that specifies $\mathcal{F}_t$-adapted processes for incentives and experimentation, $(\beta_t)_{t\geq0}$ and $(\eta_t)_{t\geq0}$, I denote with $\tilde{\xi}(\mathcal{C},\phi)$ the information rent implied by the contract $\mathcal{C}$ when beliefs are ϕ, that is,

$$\tilde{\xi}(\mathcal{C},\phi) = \mathrm{E}\left[\int_0^\infty e^{-\delta t}\eta_t\beta_t\phi_t(1-\phi_t)\,dt\,\middle|\,\mathcal{F}_0\right].$$

I will often consider the information rent implied by an optimal contract for the dual problem $\mathcal{C}^D_{V,Y,\phi}$. To simplify notation I denote this information rent as $\hat{\xi}(V,Y,\phi) = \tilde{\xi}(\mathcal{C}^D_{V,Y,\phi},\phi)$.

Throughout this appendix, I use the following notation for the left and right derivatives of a function. Consider a function of n variables $h(x_1,\ldots,x_n)$. I denote the left derivative with respect to x_i as

$$h_{x_i-}(x_1,\ldots,x_n) \equiv \lim_{\varepsilon\to0^+}\frac{h(x_1,\ldots,x_i,\ldots,x_n)-h(x_1,\ldots,x_i-\varepsilon,\ldots,x_n)}{\varepsilon},$$

and I denote the right derivative with respect to x_i as

$$h_{x_i+}(x_1,\ldots,x_n) \equiv \lim_{\varepsilon\to0^+}\frac{h(x_1,\ldots,x_i+\varepsilon,\ldots,x_n)-h(x_1,\ldots,x_i,\ldots,x_n)}{\varepsilon}.$$

Whenever the left and right derivative coincide, I use the usual notation $h_{x_i}(x_1,\ldots,x_n)$ to denote the derivative of h with respect to x_i.

Lemma 1.7. *Consider an incentive compatible contract* $\mathcal{C}$.

If $\mathrm{E}\left[\int_0^\infty e^{-rt}\left(\mu(\eta_t,\phi_t,0)\frac{\eta_t\xi_t}{\lambda\sigma}C_t^\rho\right)\,dt\Big|\mathcal{F}_0\right]$ *is finite, then*

$$\mathrm{E}\left[\int_0^\infty e^{-rt}\left(Y_t\eta_t\beta_t\phi_t(1-\phi_t)\right)\,dt\Big|\mathcal{F}_0\right]$$

$$= Y_0\tilde{\xi}(\mathcal{C},\phi_0) - \mathrm{E}\left[\int_0^\infty e^{-rt}\left(\mu(\eta_t,\phi_t,0)\frac{\eta_t\xi_t}{\lambda\sigma}C_t^\rho\right)\,dt\Big|\mathcal{F}_0\right]$$

Proof. Let

$$\tilde{Y}_t = \left[-\left(\int_0^t e^{s(\delta-r)}\mu(\eta_t,\phi_t,0)\frac{\eta_t}{\lambda\sigma}C_t^\rho\,ds\right)+\bar{Y}\right]$$

so that

$$Y_t = e^{t(r-\delta)}\tilde{Y}_t$$

For a finite $T > 0$, we can use integration by parts to obtain

$$\mathrm{E}\left[\int_0^T e^{-rt}\left(Y_t \eta_t \beta_t \phi_t (1 - \phi_t)\right) dt \Big| \mathcal{F}_0\right]$$

$$= \mathrm{E}\left[\int_0^T e^{-\delta t}\left(\tilde{Y}_t \eta_t \beta_t \phi_t (1 - \phi_t)\right) dt \Big| \mathcal{F}_0\right]$$

$$= -\mathrm{E}\Bigg\{\left[\tilde{Y}_t \int_t^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds\right]_0^T$$

$$- \int_0^T \left(\int_t^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds\right) d\tilde{Y}_t \Big| \mathcal{F}_0\Bigg\}$$

$$= \tilde{Y}_0 \mathrm{E}\left[\int_0^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds \Big| \mathcal{F}_0\right]$$

$$+ \mathrm{E}\left[\int_0^T e^{-\delta t} \left(\int_t^T e^{-\delta(s-t)} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds\right) d\tilde{Y}_t \Big| \mathcal{F}_0\right]$$

$$= \tilde{Y}_0 \mathrm{E}\left[\int_0^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds \Big| \mathcal{F}_0\right]$$

$$- \mathrm{E}\left[\int_0^T e^{-rt} \left(\int_t^T e^{-\delta(s-t)} \eta_s \beta_s \phi_s (1 - \phi_s)\, ds\right) \mu(\eta_t, \phi_t, 0) \frac{\eta_t}{\lambda \sigma} C_t^\rho \Big| \mathcal{F}_0\right]$$

Using the monotone convergence theorem and the law of iterated expectations we can conclude that

$$\mathrm{E}\left[\int_0^\infty e^{-rt}\left(Y_t \eta_t \beta_t \phi_t (1 - \phi_t)\right) dt \Big| \mathcal{F}_0\right]$$

$$= \lim_{t \to \infty} \mathrm{E}\left[\int_0^T e^{-rt}\left(Y_t \eta_t \beta_t \phi_t (1 - \phi_t)\right) dt \Big| \mathcal{F}_0\right]$$

$$= \tilde{Y}_0 \tilde{\xi}(\mathcal{C}, \phi_0) - \mathrm{E}\left[\int_0^\infty e^{-rt}\left(\mu(\eta_t, \phi_t, 0)\frac{\eta_t \xi_t}{\lambda \sigma} C_t^\rho\right) dt \Big| \mathcal{F}_0\right].$$

103

Since $\tilde{Y}_0 = Y_0$, this concludes the proof. $\qquad\square$

Lemma 1.8. $G^*(V, Y, \phi)$ *is decreasing and concave in Y.*

Proof. Fix V and ϕ, consider Y^0 and Y^1 such that $Y^0 \leq Y^1$. Then

$$J^*(V, Y^1, \phi) \leq J(V, Y^1, \phi | \mathcal{C}^D_{V,Y^0,\phi}) = J^*(V, Y^0, \phi) + (Y^0 - Y^1)\xi_0 \leq J^*(V, Y^0, \phi),$$

where the first equality follows by applying Lemma 1.7.

To show convexity, consider $\nu \in [0, 1]$ and define $Y^\nu = \nu Y^0 + (1 - \nu)Y^1$. Then

$$
\begin{aligned}
&\nu J^*(V, Y^0, \phi) + (1 - \nu)J^*(V, Y^1, \phi) \\
&\leq \nu J(V, Y^0, \phi | \mathcal{C}^D_{V,Y^\nu,\phi}) + (1 - \nu)J(V, Y^1, \phi | \mathcal{C}^D_{V,Y^\nu,\phi}) \\
&= J^*(V, Y^\nu, \phi)
\end{aligned}
$$

The last equality holds because, under contract $\mathcal{C}^D_{V,Y^\nu,\phi} = ((C^\nu_t, K^\nu_t, \eta^\nu_t, 0))_{t \geq 0}$,

$$Y^\nu_t = e^{(r-\delta)t}\left[-\left(\int_0^t e^{(\delta-r)s}(\alpha + \sigma\eta^\nu_t \phi^\nu_t)\frac{\eta^\nu_t}{\lambda\sigma}(C^\nu_t)^\rho \, ds \right) + Y^\nu \right].$$

$\qquad\square$

Lemma 1.9 (Weak Duality). *Let $\mathcal{C}$ be an incentive compatible contract that delivers expected lifetime utility V to the agent and that implies an information rent ξ. Then*

$$J(V, \xi, \phi | \mathcal{C}) = G(V, Y, \phi | \mathcal{C}) + Y\xi$$

104

Proof. This follows immediately from Lemma (1.7), since

$$J(V, \xi, \phi | \mathcal{C}) = \mathrm{E}\left[\int_0^\infty e^{-rt}\left(\frac{C_t}{q(\eta_t)} - \mu(\eta_t, \phi_t, 0)\frac{\beta_t - \eta_t\xi_t}{\lambda\sigma}C_t^\rho\right) dt \,\bigg|\, \mathcal{F}_0\right]$$

$$= G(V, Y, \phi | \mathcal{C}) + \mathrm{E}\left[\int_0^\infty e^{-rt}\left(\mu(\eta_t, \phi_t, 0)\frac{\eta_t\xi_t}{\lambda\sigma}C_t^\rho\right) dt \,\bigg|\, \mathcal{F}_0\right]$$

$$+ \mathrm{E}\left[\int_0^\infty e^{-rt}\left(Y_t\eta_t\beta_t\phi_t(1 - \phi_t)\right) dt \,\bigg|\, \mathcal{F}_0\right]$$

$$= G(V, Y, \phi | \mathcal{C}) + Y\xi.$$

$\square$

Lemma 1.10. *The dual cost function is homogeneous with $G^*(V, Y, \phi) = \hat{v}g^*(y, \phi)$ for a continuous function $g^*(y, \phi)$.*

Proof. Let

$$\hat{v} = ((1 - \rho)V)^{\frac{1}{1-\rho}}$$

be the consumption equivalent of the manager's continuation value and let

$$y = (1 - \phi)\hat{v}^{-\rho}Y$$

be scaled version of multiplier Y. Define the scaled control variables

$$c_t = \frac{C_t}{\hat{v}_t} \quad k_t = \frac{K_t}{\hat{v}_t} \quad \hat{\beta}_t = \frac{\beta_t}{(1 - \rho)V_t}. \tag{1.41}$$

The laws of motion of promised value $\hat{v}$ and multiplier y can be obtained by applying

Ito's lemma,

$$\frac{d\hat{v}_t}{\hat{v}_t} = \left(\frac{\delta}{1-\rho} - \frac{c_t^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_t^2\right)dt + \hat{\beta}_t\hat{v}_t dW_t^{\mathbb{C}} \tag{1.42}$$

and

$$dy_t = -(1-\phi_t)\mu(\eta_t, \phi_t, 0)\frac{\eta_t}{\lambda\sigma}c_t^{\rho}dt$$

$$+ y_t\left(r - \frac{\delta}{1-\rho} + \rho\frac{c_t^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_t^2 + \rho\eta_t\hat{\beta}_t\phi_t\right)dt - y_t[\rho\hat{\beta}_t + \eta_t\phi_t]dW_t^{\mathbb{C}}. \tag{1.43}$$

From equation (1.42), we obtain

$$\hat{v}_t = \hat{v}_0 \exp\left\{\int_0^t\left(\frac{\delta}{1-\rho} - \frac{c_s^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_s^2 - \frac{1}{2}\hat{\beta}^2\right)ds + \int_0^t\hat{\beta}_s dW_s^{\mathbb{C}}\right\} \tag{1.44}$$

Using expressions (1.41) and equation (1.44), we can write the objective function as

$$\hat{v}_0\mathrm{E}\left[\int_0^\tau e^{-\int_0^t r - \left(\frac{\delta}{1-\rho} - \frac{c_s^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_s^2\right)ds} B_t\left(\frac{c_t}{q(\eta)} - \mu(\eta_t, \phi_t, 0)\frac{\hat{\beta}_t}{\lambda\sigma}c_t^{\rho} - y_t\phi_t\eta_t\hat{\beta}_t\right)dt\,\Big|\,\mathcal{F}_0\right] \tag{1.45}$$

where B_t is a density process,

$$B_t = \exp\left\{\int_0^t\hat{\beta}_s dW_s^{\mathbb{C}} - \frac{1}{2}\int_0^t\hat{\beta}^2\,ds\right\}.$$

Minimizing (1.45) is equivalent to minimizing the expectation in (1.45) subject to beliefs and the law of motion of multiplier y, (1.43). To see why, notice that any policy that is optimal for initial states $\hat{v}_0, y_0,$ and ϕ_0 must also be optimal for $\hat{v}_0', y_0,$

106

and ϕ_0, with $\hat{v}'_0 \neq \hat{v}_0$.

Le $g^*(y, \phi)$ be defined by

$$g^*(y_0, \phi_0) =$$

$$\inf_{(c_t, \hat{\beta}_t, \eta_t)_{t \geq 0}} \tilde{\mathbb{E}} \left[\int_0^\tau e^{-\int_0^t r - \left(\frac{\delta}{1-\rho} - \frac{c_s^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}_s^2 \right) ds} \left(\frac{c_t}{q(\eta)} - \mu(\eta_t, \phi_t, 0)\frac{\hat{\beta}_t}{\lambda\sigma}c_t^\rho \right. \right.$$

$$\left. \left. - y_t\phi_t\eta_t\beta_t \right) dt \middle| \mathcal{F}_0 \right]$$

$$\text{s.t.} \quad d\phi_t = \eta_t\phi_t(1 - \phi_t)\hat{\beta}dt + \eta_t\phi_t(1 - \phi_t)d\tilde{W}_t^{\mathcal{C}} \tag{1.46}$$

$$dy_t = -(1 - \phi_t)\mu(\eta_t, \phi_t, 0)\frac{\eta_t}{\lambda\sigma}c_t^\rho dt$$

$$+ y_t \left(r - \frac{\delta}{1 - \rho} + \rho\frac{c_t^{1-\rho}}{1 - \rho} + \frac{1}{2}\rho\hat{\beta}_t^2 + \rho\eta_t\hat{\beta}_t\phi_t \right) dt$$

$$- y_t[\rho\hat{\beta}_t + \eta_t\phi_t]\hat{\beta}dt - y_t[\rho\hat{\beta}_t + \eta_t\phi_t]d\tilde{W}_t^{\mathcal{C}},$$

where $\tilde{W}_t^{\mathcal{C}}$ is a Brownian motion under the measure $\tilde{Q}$ such that $\frac{d\tilde{Q}}{dP}|_{\mathcal{F}_t} = B_t$. Then $G^*(V, Y, \phi) = \hat{v}g^*(y, \phi)$ where $\hat{v} = ((1 - \rho)V)^{\frac{1}{1-\rho}}$ and $y = (1 - \phi)\hat{v}^{-\rho}Y$. $\qquad \square$

Lemma 1.11. *If $Y < 0$, G^* is differentiable with respect to Y.*

Proof. Let us use Lemma 1.10 to obtain $G^*(V, Y, \phi) = \hat{v}g^*(y, \phi)$ for $y = (1-\phi)\hat{v}^{-\rho}Y$. Since G^* is concave in Y, then g^* is concave in y. Therefore, the right and left derivatives of $\hat{v}g^*$ with respect to y always exist in the interior of the domain. Moreover, $G_Y^*(V, Y, \phi)$ exists if and only if $\hat{v}g_y^*(y, \phi)$ exists. Suppose that there exist a point, $(\hat{v}_0, y_0, \phi_0)$ where $g_{y+}^* < g_{y-}^*$. Consider a test function F such that $F(\hat{v}, y, \phi) - \hat{v}g^*(y, \phi)$ has a local minimum at $(\hat{v}_0, y_0, \phi_0)$ and such that $F(\hat{v}_0, y_0, \phi_0) -$

$\hat{v}_0 g^*(y_0, \phi_0) = 0.$

Then consider a $p \in (g^*_{y+}, g^*_{y-})$ and take another test function

$$F_\varepsilon(\hat{v}, y, \phi) = F(\hat{v}, y_0, \phi) + \hat{v}p(y - y_0) - \frac{1}{2\epsilon}\hat{v}(y - y_0)^2.$$

For any arbitrary $\varepsilon > 0$, $F_\varepsilon(\hat{v}, y, \phi) \geq \hat{v}g^*(y, \phi)$ in a neighborhood of $(\hat{v}_0, y_0, \phi_0)$ and $(\hat{v}_0, y_0, \phi_0)$ is a minimizer of $F_\varepsilon(\hat{v}, y, \phi) - \hat{v}g^*(y, \phi)$. Then the viscosity subsolution property of F_ε (Pham, 2009) implies that

$$
\begin{aligned}
rg^*(y, \phi) - \inf_{c, \hat{\beta}, \eta} \Bigg\{ & \frac{c}{q(e)} - (\alpha + \sigma\eta\phi)\hat{\beta}\frac{c^\rho}{\sigma\lambda} - y\hat{\beta}\eta\phi \\
& + g^*(y, \phi)\left(\frac{\delta}{1 - \rho} - \frac{c^{1-\rho}}{1 - \rho} + \frac{1}{2}\rho\hat{\beta}^2\right) \\
& + p\Bigg[-(1 - \phi_t)\mu(\eta, \phi, 0)\frac{\eta}{\lambda\sigma}c^\rho \\
& \qquad\qquad + y\left(r - \frac{\delta}{1 - \rho} + \rho\frac{c^{1-\rho}}{1 - \rho} + \frac{1}{2}\rho\hat{\beta}^2 + \rho\eta\hat{\beta}\phi\right)\Bigg] \quad (1.47) \\
& - py[\rho\hat{\beta} + \eta\phi]\hat{\beta} + F_\phi(y, \phi)\eta\phi(1 - \phi)\hat{\beta} \\
& \frac{1}{2}F_{\phi\phi}(y, \phi)\eta^2\phi^2(1 - \phi)^2 \\
& - \frac{1}{\varepsilon}[\rho\hat{\beta} + \eta\phi]^2 \Bigg\} \leq 0.
\end{aligned}
$$

$[\rho\hat{\beta} + \eta\phi]^2$ is always strictly positive. To see why, suppose that $\eta = 0$. In this case, since $g^*(y, \phi)$ is decreasing and concave in y, we must have that $g^*(y, \phi) - \frac{1}{2}py \geq 0$. Then the first order condition with respect to $\hat{\beta}$ holds and imply that

$$\alpha\frac{c^\rho}{\lambda\sigma} = (g^*(y, \phi) - \frac{1}{2}py)\hat{\beta},$$

which in turn implies that $\hat{\beta} > 0$ because c is also positive, since $g^*(y, \phi) - \rho p y \geq 0$ and the marginal utility of consumption is infinite at $c = 0$. Since $[\rho \hat{\beta} + \eta \phi]^2$ is strictly positive and ε is arbitrary, the inequality in (1.47) is a contradiction. Hence, $\hat{v} g^*_{y+}(y, \phi) = \hat{v} g^*_{y-}(y, \phi)$, from which it immediately follows that $G^*_{Y+}(V, Y, \phi) = G^*_{Y-}(V, Y, \phi)$. $\qquad \square$

Lemma 1.12. *If $Y < 0$, then*

$$G^*_Y(V, Y, \phi) = -\hat{\xi}(V, Y, \phi).$$

Proof. Consider the left derivative

$$G^*_{Y-}(V, Y, \phi) = \lim_{\epsilon \to 0^+} \frac{G^*(V, Y, \phi) - G^*(V, Y - \epsilon, \phi)}{\epsilon}.$$

Since $G^*(V, Y - \epsilon, \phi) < G(V, Y - \epsilon, \phi | \mathcal{C}^D_{V,Y,\phi})$, then

$$G^*_{Y-}(V, Y, \phi) \geq \lim_{\epsilon \to 0^+} \frac{G^*(V, Y, \phi) - G(V, Y - \epsilon, \phi | \mathcal{C}^D_{V,Y,\phi})}{\epsilon} =$$

$$= -\mathrm{E}\left[\int_0^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s) \, ds \Big| \mathcal{F}_0\right] = -\hat{\xi}(V, Y, \phi)$$

Similarly, consider the right derivative

$$G^*_{Y+}(V, Y, \phi) = \lim_{\epsilon \to 0^+} \frac{G^*(V, Y + \epsilon, \phi) - G^*(V, Y, \phi)}{\epsilon}.$$

Then

$$G^*_{Y+}(V, Y, \phi) \leq \lim_{\epsilon \to 0^+} \frac{G(V, Y + \epsilon, \phi | \mathcal{C}^D_{V,Y,\phi}) - G^*(V, Y, \phi)}{\epsilon}$$

$$= -\mathrm{E}\left[\int_0^T e^{-\delta s} \eta_s \beta_s \phi_s (1 - \phi_s) \, ds \Big| \mathcal{F}_0\right] = -\hat{\xi}(V, Y, \phi)$$

Hence

$$G^*_{Y+}(V, Y, \phi) \leq -\hat{\xi}(V, Y, \phi) \leq G^*_{Y-}(V, Y, \phi).$$

Since G^* is differentiable with respect to Y when $Y < 0$, we conclude that $G^*_Y(V, Y, \phi) = -\hat{\xi}(V, Y, \phi)$. $\square$

Lemma 1.13. $\hat{\xi}(V, 0, \phi) = \lim_{\epsilon \to 0^+} \hat{\xi}(V, -\epsilon, \phi)$

Proof. Let $\xi_0 = \hat{\xi}(V_0, 0, \phi_0)$. Since $Y_t < 0$ for all $t > 0$, this means that $\xi_t = \hat{\xi}(V_t, Y_t, \phi_t) \leq \lim_{\epsilon \to 0^+} \hat{\xi}(V, -\epsilon, \phi)$. Define $\xi^{-\epsilon} \equiv \lim_{\epsilon \to 0^+} \hat{\xi}(V, -\epsilon, \phi)$. Suppose that $D \equiv \hat{\xi}(V, 0, \phi) - \xi^{-\epsilon} > 0$.

In general,

$$\xi_0 = \mathrm{E}\left[\int_0^t e^{-\delta s} \beta_s \eta_s \phi_s (1 - \phi_s) \, ds \Big| \mathcal{F}_0\right] + \mathrm{E}\left[e^{-Dt} \xi_t | \mathcal{F}_t\right]$$

$$\leq \mathrm{E}\left[\int_0^t e^{-\delta s} \beta_s \eta_s \phi_s (1 - \phi_s) \, ds \Big| \mathcal{F}_0\right] + \xi^{-\epsilon},$$

which implies that

$$D \leq \mathrm{E}\left[\int_0^t e^{-\delta s} \beta_s \eta_s \phi_s (1 - \phi_s) \, ds \Big| \mathcal{F}_0\right].$$

110

Since t is arbitrary, we conclude that $D = 0$. □

Lemma 1.14. *For any arbitrary ξ and Y,*

$$G^*(V, Y, \phi) \leq J^*(V, \xi, \phi) - Y\xi.$$

Proof. Using Lemma 1.9, we can see that, for any arbitrary ξ and Y,

$$J^*(V, \xi, \phi) - Y\xi = J(V, \xi, \phi | \mathcal{C}^P_{V,\xi,\phi}) - Y\xi = G(V, Y, \phi | \mathcal{C}^P_{V,\xi,\phi}) \geq G^*(V, Y, \phi).$$

□

Lemma 1.15. *For any arbitrary Y,*

$$G^*(V, Y, \phi) \geq J^*(V, \hat{\xi}(V, Y, \phi), \phi) - Y\hat{\xi}(V, Y, \phi).$$

Proof. Using Lemma 1.9, we can see that, for any arbitrary Y,

$$J^*(V, \hat{\xi}(V, Y, \phi), \phi) \leq J(V, \hat{\xi}(V, Y, \phi), \phi | \mathcal{C}^D_{V,Y,\phi}) = G^*(V, Y, \phi) + Y\hat{\xi}(V, Y, \phi).$$

□

Lemma 1.16. *If $Y = \arg\sup_{Y' \leq 0}\{G^*(V, Y', \phi) + Y'\xi\}$, then*

$$\hat{\xi}(V, Y, \phi) = \xi, \quad \text{if } Y < 0$$

and

$$\hat{\xi}(V, 0, \phi) \leq \xi.$$

Proof. Since $G^*(V, Y', \phi)$ is concave and differentiable with respect to Y for $Y < 0$, the first-order condition holds and

$$-G_Y^*(V, Y, \phi) = \xi$$

if $Y < 0$, and

$$-G_{Y^-}^*(V, 0, \phi) \leq \xi$$

if $Y = 0$. Using Lemmas 1.12 and 1.13, we conclude the proof. $\qquad\square$

Lemma 1.17. $J^*(V, \hat{\xi}(V, 0, \phi), \phi) \leq J^*(V, \xi, \phi)$ *for all* $\xi \geq \hat{\xi}(V, 0, \phi)$.

Proof. If $\xi \geq \hat{\xi}(V, 0, \phi)$ then $0 = \arg\sup_{Y' \leq 0}\{G^*(V, Y', \phi) + Y'\xi\}$. Let

$$\bar{\xi} = \arg\min_{\xi \geq \hat{\xi}(V, 0, \phi)}\{J^*(V, \xi, \phi)\}.$$

From Lemma 1.14, it follows that $G^*(V, 0, \phi) \leq J(V, \bar{\xi}, \phi)$. But, by Lemma 1.15, we must also have $G^*(V, 0, \phi) \geq J^*(V, \xi(V, 0, \phi), \phi)$. Therefore, $J(V, \xi, \phi)$ is minimized by $\xi = \xi(V, 0, \phi)$.

$\qquad\square$

Lemma 1.18. *Consider* $\xi \leq \hat{\xi}(V, 0, \phi)$ *and let* $Y = \arg\sup_{Y' \leq 0}\{G^*(V, Y', \phi) + Y'\xi\}$. *Then* $\mathcal{C}_{V, Y, \phi}^D$ *is optimal also for the primal problem for initial states* V, ξ *and* ϕ.

Moreover,

$$J^*(V, \xi, \phi) = \sup_{Y' \leq 0} \{G^*(V, Y', \phi) + Y'\xi\}. \tag{1.48}$$

Proof. By Lemma 1.16, $\mathcal{C}_{V,Y,\phi}^D$ implies information rent ξ, so it is a candidate for an optimal contract given states V, ξ and ϕ. Then the following holds:

$$J(V, \xi, \phi | \mathcal{C}_{V,Y,\phi}^D) \geq J^*(V, \xi, \phi)$$

$$\geq \sup_{Y' \leq 0} \{G^*(V, Y', \phi) + Y'\xi\} = G^*(V, Y, \phi) + Y\xi$$

$$= G^*(V, Y, \phi) + Y\hat{\xi}(V, Y, \phi).$$

The first inequality holds because $\mathcal{C}_{V,Y,\phi}^D$ cannot be better than the optimal contract, the second inequality holds because of Lemma 1.14. The following equality holds because $Y = \arg\sup_{Y' \leq 0}\{G^*(V, Y', \phi) + Y'\xi\}$. Finally, the last equality holds because of Lemma 1.16.

By Lemma 1.9, $J(V, \xi, \phi | \mathcal{C}_{V,Y,\phi}^D) = G^*(V, Y, \phi) + Y\hat{\xi}(V, Y, \phi)$. Hence, all the inequalities hold with equality, $\mathcal{C}_{V,Y,\phi}^D$ is optimal for the primal problem and

$$J^*(V, \xi, \phi) = \sup_{Y' \leq 0} \{G^*(V, Y', \phi) + Y'\xi\}.$$

$\square$

Lemma 1.19. $J^*(V, \xi, \phi)$ *is differentiable, decreasing and convex in ξ for $\xi \leq \hat{\xi}(V, 0, \phi)$.*

Proof. When $\xi \leq \hat{\xi}(V, 0, \phi)$, the solution to the maximization problem in (1.48)

113

is interior and the envelope theorem holds (Milgrom and Segal, 2002). Therefore $J_\xi^*(V, \xi, \phi)$ exists and it is equal to Y, where $Y = \arg\sup_{Y' \leq 0}\{G^*(V, Y', \phi) + Y'\xi\}$. Since $Y \leq 0$, $J^*(V, \xi, \phi)$ is decreasing in ξ for $\xi \leq \hat{\xi}(V, 0, \phi)$.

To prove convexity consider $\xi^0 \leq \hat{\xi}(V, 0, \phi)$ and $\xi^1 \leq \hat{\xi}(V, 0, \phi)$ and let $\xi^\nu = (1 - \nu)\xi^0 + \nu\xi^1$. Using the concavity of $G^*(V, Y, 0)$ and Lemma 1.18 we obtain

$$
\begin{aligned}
J^*(V, \xi^\nu, \phi) &= \sup_{Y' \leq 0}\left\{G^*(V, Y', \phi) + Y'\xi^\nu\right\} \\
&\leq (1 - \nu)\sup_{Y' \leq 0}\left\{G^*(V, Y', \phi) + Y'\xi^0\right\} + \nu\sup_{Y' \leq 0}\left\{G^*(V, Y', \phi) + Y'\xi^1\right\} \\
&= (1 - \nu)J^*(V, \xi^0, \phi) + \nu J^*(V, \xi^1, \phi).
\end{aligned}
$$

$\square$

Lemma 1.20. $\hat{\xi}(V, 0, \phi)$ *is a global minimum of* $J^*(V, \xi, \phi)$ *with respect to* ξ.

Proof. From Lemma 1.17, $\hat{\xi}(V, 0, \phi)$ is a global minimum for $\xi \geq \hat{\xi}(V, 0, \phi)$. From Lemma 1.19, $\hat{\xi}(V, 0, \phi)$ is a global minimum also for $\xi \leq \hat{\xi}(V, 0, \phi)$. Therefore, $\xi \geq \hat{\xi}(V, 0, \phi)$ is a global minimum $J^*(V, \xi, \phi)$ with respect to ξ. $\square$

1.D.1   Proof of Propositions 1.6 and 1.7

Proof. Consider an initial promised value for the agent, V_0, and initial beliefs ϕ_0. Since $\hat{\xi}(V_0, 0, \phi_0)$ is a global minimum for $J^*(V_0, \xi, \phi_0)$ with respect to ξ (by Lemma 1.20), the principal sets $\xi_0 = \hat{\xi}(V_0, 0, \phi_0)$. By Lemma 1.18, $\mathcal{C}_{V_0, 0, \phi_0}^D$ is the optimal contract for the principal at time zero, and equation (1.48) holds. At any time t $Y_t \leq 0$, and combining Lemmas 1.16, 1.18, and 1.19, we obtain that $\xi_t = \hat{\xi}(V_t, Y_t, \phi_t) = $

114

$-G^*_Y(V_t, Y_t, \phi_t)$ and that $Y_t = J^*_\xi(V_t, \xi_t, \phi_t)$. We then obtain Proposition 1.7.

Let $\bar{\xi}(V_t, \phi_t) = \hat{\xi}(V_t, 0, \phi_t)$. Considering that $J^*_\xi(V_t, \xi_t, \phi_t) = Y_t \leq 0$ and that $\xi_t = -G^*_Y(V_t, Y_t, \phi_t) \leq -G^*_{Y^-}(V_t, 0, \phi_t) \leq \hat{\xi}(V_t, 0, \phi_t)$ (by the concavity of G^* and by Lemma 1.13), we obtain Proposition 1.6.

□

Appendix 1.E   Proofs for Appendix 1.A

1.E.1   Proof of Lemma 1.4

Proof. The first part of the Lemma follows directly from the definition of Markovian contract and weakly renegotiation-proof contract.

Suppose that the optimal renegotiation-proof contract is unique. Consider stopping times $t \geq 0$ and $t' \geq 0$ such that $V_t = V_{t'}$ and $\phi_t = \phi_{t'}$. By uniqueness of the contract, it follows that $\mathcal{O}(\mathcal{C}, t) = \mathcal{O}(\mathcal{C}, t')$. To see why, if $\mathcal{O}(\mathcal{C}, t) \neq \mathcal{O}(\mathcal{C}, t')$ then contract $\mathcal{C}'$ such that $\mathcal{O}(\mathcal{C}', s) = \mathcal{O}(\mathcal{C}, s)$ for all $s \neq t$ and $\mathcal{O}(\mathcal{C}', t) = \mathcal{O}(\mathcal{C}, t')$ would be incentive-compatible given (V_0, ϕ_0), weakly renegotiation-proof and payoff-equivalent to $\mathcal{C}$, thus contradicting the uniqueness of the optimal renegotiation-proof contract.

□

1.E.2 Proof of Proposition 1.8

Proof. Using the expression (1.11) for ξ_t, we obtain

$$
\begin{aligned}
\xi_t &= \mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}\beta_s\eta_s\phi_s(1-\phi_s)\,ds\,\Big|\,\mathcal{F}_t\right] \\
&= \mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}(1-\rho)V_s\hat{\beta}_s\eta_s\phi_s(1-\phi_s)\,ds\,\Big|\,\mathcal{F}_t\right] \\
&= \mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}\hat{v}_s^{1-\rho}\hat{\beta}_s\eta_s\phi_s(1-\phi_s)\,ds\,\Big|\,\mathcal{F}_t\right] \\
&= \hat{v}_0^{1-\rho}\mathrm{E}\left[\int_t^\infty e^{-\delta(s-t)}\left(\frac{\hat{v}_s}{\hat{v}_t}\right)^{1-\rho}\hat{\beta}_s\eta_s\phi_s(1-\phi_s)\,ds\,\Big|\,\mathcal{F}_t\right] \\
&= (1-\rho)V_t z_t,
\end{aligned}
$$

where

$$
z_t = \mathrm{E}\left[\int_t^\infty e^{\int_t^s\left(-c_u^{1-\rho}\right)du}\tilde{B}_{ts}\hat{\beta}_s\eta_s\phi_s(1-\phi_s)\,ds\,\Big|\,\mathcal{F}_t\right], \tag{1.49}
$$

for a density process $\tilde{B}_{ts}$ such that

$$
\tilde{B}_{ts} = \exp\left\{\int_t^s(1-\rho)\hat{\beta}_u dW_u^{\mathcal{C}} - \frac{1}{2}\int_t^s(1-\rho)^2\hat{\beta}_u^2\,du\right\}.
$$

The objective function can be written as

$$
\hat{v}_0\mathrm{E}\left[\int_0^\tau e^{-\int_0^t r-\left(\frac{\delta}{1-\rho}-\frac{c_s^{1-\rho}}{1-\rho}+\frac{1}{2}\rho\hat{\beta}_s^2\right)ds}B_t\left(\frac{c_t}{q(\eta)}-\mu(\eta_t,\phi_t,0)\frac{\hat{\beta}_t-\eta_t z_t}{\lambda\sigma}c_t^\rho\right)dt\,\Big|\,\mathcal{F}_0\right], \tag{1.50}
$$

116

where B_t is a density process,

$$B_t = \exp\left\{ \int_0^t \hat{\beta}_s dW_s^{\mathcal{C}} - \frac{1}{2} \int_0^t \hat{\beta}^2\, ds \right\}.$$

Minimizing equation (1.50) is equivalent to minimizing the expectation in (1.50) subject to the law of motion of beliefs (under the probability measure implied by the density process B_t) and where z_t is given by (1.49). To see why, notice that if a strategy $((c_t)_{t\geq 0}, (\hat{\beta}_t)_{t\geq 0}, (\eta_t)_{t\geq 0})$ is optimal for $\hat{v}_0$ and ϕ, then it must also be optimal for $\hat{v}_0'$ and ϕ, even if $\hat{v}_0' \neq \hat{v}_0$. Equation (1.27) in Proposition 1.8 is the associated HJB equation.

Because the strategy $((c_t)_{t\geq 0}, (\hat{\beta}_t)_{t\geq 0}, (\eta_t)_{t\geq 0})$ depends only on beliefs ϕ, then the information rent implied by the contract is given by $(1 - \rho)V_t z_R(\phi_t)$ where

$$z_R(\phi_t) = \mathrm{E}\left[\int_t^\infty e^{\int_t^s (-c(\phi_u)^{1-\rho})\, du} \tilde{B}_{ts} \hat{\beta}(\phi_s) \eta(\phi_s) \phi_s (1 - \phi_s)\, ds \,\middle|\, \mathcal{F}_t \right], \qquad (1.51)$$

Equation (1.28) in Proposition 1.8 is therefore obtained in the following way. According to equation (1.51), the drift of $z_R(\phi_t)$ (under the probability measure implied by the density process $\tilde{B}_t$) is

$$c_R(\phi_t)^{1-\rho} z_R(\phi_t) - \hat{\beta}_R(\phi_t) \eta_R(\phi_t) \phi_t (1 - \phi_t).$$

The drift of $z_R(\phi_t)$ (under the probability measure implied by the density process

117

$\tilde{B}_t$) can also be obtained by using Ito's lemma:

$$(1 - \rho)\hat{\beta}_R(\phi)\eta_R(\phi)\phi(1 - \phi)z_R'(\phi) + \frac{1}{2}\eta_R(\phi)^2\phi^2(1 - \phi)^2 z_R''(\phi).$$

Equating the two drifts, we obtain the differential equation (1.28) that characterizes z_R.

$\square$

1.E.3 Proof of Proposition 1.9

Proof. The proof is analogous to the proof of Proposition 1.8 and it is therefore omitted. We just need to use A_t, instead of $\hat{v}_t$, as the scaling process. We also need to consider that

$$\beta_t = V_A(A_t, \phi_t)\Theta_t\sigma + V_\phi(A_t, \phi_t)\eta_t\phi_t(1 - \phi_t)$$

when deriving the differential equation for the information rent.

$\square$

1.E.4 Proof of Proposition 1.10

Proof. Consider the HJB equation (1.30) in Proposition 1.9. I will show that $v_M(\phi) = J_R(\phi)^{-(1-\rho)}$ with $c_M(\phi) = c_R(\phi)J_R(\phi)$, $\eta_M(\phi) = \eta_R(\phi)$, and $\hat{\Theta}(\phi)$ such that $v_M(\phi)\hat{\Theta}\sigma + \frac{v_M'(\phi)}{1-\rho}\eta(\phi)\phi(1 - \phi) = \hat{\beta}_R(\phi)J_R(\phi)^{-(1-\rho)}$ is a solution to that HJB equation (1.30) with $z_M(\phi) = \left(\frac{J_R(\phi)^{1-\rho}}{1-\rho}\right)^{-1}z_R(\phi)$.

118

Using $v_M(\phi) = J_R(\phi)^{-(1-\rho)}$ in the HJB equation (1.30), we obtain

$$\frac{\delta}{1-\rho} = \max_{c,\Theta,\eta,k} \left\{ \frac{c^{1-\rho}}{1-\rho} J_R(\phi)^{1-\rho} + \left[r - \frac{c}{q(\eta)} + \mu(\eta,\phi,0)k - \frac{1}{2}\rho\hat{\Theta}^2\sigma^2 \right] \right.$$

$$- (1-\rho)\eta\hat{\Theta}\sigma\phi(1-\phi)J_R(\phi)^{-1}J_R'(\phi)$$

$$\left. - \frac{1}{2}(\rho-2)\eta^2\phi^2(1-\phi)^2 J_R(\phi)^{-2}J_R'(\phi)^2 - \frac{1}{2}\eta^2\phi^2(1-\phi)^2 J_R(\phi)^{-1}J_R''(\phi) \right\}$$

s.t. $(1-\rho)v_M(\phi)\hat{\Theta}\sigma + v_M'(\phi)\eta_t\phi(1-\phi) = (1-\rho)c^{-\rho}k\lambda\sigma + \eta z_M(\phi)$.

Substituting for $\hat{\beta}(\phi)$, the incentive-compatibility constraint becomes

$$\hat{\beta} = J_R(\phi)^{1-\rho}c^{-\rho}k\lambda\sigma + \eta J_R(\phi)^{1-\rho}\frac{z_M(\phi)}{1-\rho},$$

while

$$\hat{\Theta}\sigma = \hat{\beta} + \phi(1-\phi)J_R(\phi)^{-1}J_R'(\phi). \tag{1.52}$$

After some simplifications, the HJB becomes

$$\frac{\delta}{1-\rho} = \max_{c,\hat{\beta},\eta,k} \left\{ \frac{c^{1-\rho}}{1-\rho} J_R(\phi)^{1-\rho} + \left[r - \frac{c}{q(\eta)} + \mu(\eta,\phi,0)k \right] - \frac{1}{2}\rho\hat{\beta}^2 \right.$$

$$\left. - \phi(1-\phi)\eta\hat{\beta}J_R(\phi)^{-1}J_R'(\phi) - \frac{1}{2}\phi^2(1-\phi)^2\eta^2 J_R(\phi)^{-1}J_R''(\phi) \right\}$$

s.t. $\hat{\beta} = J_R(\phi)^{1-\rho}c^{-\rho}k\lambda\sigma + \eta J_R(\phi)^{1-\rho}\frac{z_M(\phi)}{1-\rho}$.

Consider now $\tilde{c} = cJ_R(\phi)$, $\tilde{k} = J_R(\phi)$ and $z_R(\phi) = J_R(\phi)^{1-\rho}\frac{z_M(\phi)}{1-\rho}$. The HJB

can thus be written as

$$\frac{\delta}{1-\rho}J_R(\phi) = \max_{\tilde{c},\hat{\beta},\eta,\tilde{k}} \left\{ \frac{\tilde{c}^{1-\rho}}{1-\rho} + \left[r - \frac{\tilde{c}}{q(\eta)} + \mu(\eta,\phi,0)\tilde{k} \right] - \frac{1}{2}\rho\hat{\beta}^2 J_R(\phi) \right.$$

$$\left. - \phi(1-\phi)\eta\hat{\beta}J_R'(\phi) - \frac{1}{2}\phi^2(1-\phi)^2\eta^2 J_R''(\phi) \right\}$$

$$\text{s.t. } \hat{\beta} = \tilde{c}^{-\rho}\tilde{k}\lambda\sigma + \eta z_R(\phi).$$

This equation is equivalent to

$$rJ_R(\phi) = \min_{\tilde{c},\hat{\beta},e} \left\{ \frac{\tilde{c}}{q(\eta)} - \mu(\eta,\phi,0)\frac{\hat{\beta} - \eta z_R(\phi)}{\sigma\lambda}\tilde{c}^\rho + J_R(\phi)\left(\frac{\delta}{1-\rho} - \frac{\tilde{c}^{1-\rho}}{1-\rho} + \frac{1}{2}\rho\hat{\beta}^2 \right) \right.$$

$$\left. + \hat{\beta}\eta\phi(1-\phi){J_R}'(\phi) + \frac{1}{2}\eta^2\phi^2(1-\phi)^2{J_R}''(\phi) \right\}.$$

Once we verify that $z_R(\phi) = J_R(\phi)^{1-\rho}\frac{z_M(\phi)}{1-\rho}$, then the equation that we have just derived coincides with the HJB equation (1.27) in Proposition 1.8. In particular, the optimal controls are $c_R(\phi)$, $\hat{\beta}_R(\phi)$ and $\eta_R(\phi)$. It therefore remains to verify that $z_R(\phi) = J_R(\phi)^{1-\rho}\frac{z_M(\phi)}{1-\rho}$. Combining the HJB equation (1.30) with the ODE (1.32), we obtain

$$\left[-\frac{c_M(\phi)^{1-\rho}}{v_M(\phi)} - (1-\rho)\sigma\eta\hat{\Theta}\phi(1-\phi)\frac{v_M'(\phi)}{v_M(\phi)} - \frac{1}{2}\eta^2\phi^2(1-\phi)^2\frac{v_M''(\phi)}{v_M(\phi)} \right] z_M(\phi) =$$

$$(1-\rho)\eta_M(\phi)\phi(1-\phi)\hat{\Theta}(\phi)\sigma v_M(\phi) + \eta_M(\phi)^2\phi^2(1-\phi)^2 v_M'(\phi)$$

$$+ (1-\rho)\eta_M(\phi)\phi(1-\phi)\hat{\Theta}(\phi)\sigma z_M'(\phi) + \frac{1}{2}\eta_M(\phi)^2\phi^2(1-\phi)^2 z_M''(\phi).$$

We can then use substitutions and simplifications analogous to the ones we used for the HJB equation (1.27), and obtain

$$z_R(\phi)c_R(\phi)^{1-\rho} - \hat{\beta}_R(\phi)\eta_R(\phi)\phi(1-\phi) - (1-\rho)\hat{\beta}_R(\phi)\eta_R(\phi)\phi(1-\phi)z_R'(\phi)$$
$$= \frac{1}{2}\eta_R(\phi)^2\phi^2(1-\phi)^2 z_R''(\phi).$$

We have therefore verified that $v_M(\phi) = J_R(\phi)^{-(1-\rho)}$, $c_M(\phi) = c_R(\phi)J_R(\phi)$, $\eta_M(\phi) = \eta_R(\phi)$, $\hat{\Theta}_M(\phi)\sigma = \hat{\beta}_T(\phi) + \phi(1-\phi)J_R(\phi)^{-1}J_R'(\phi)$, and $z_M(\phi) = \left(\frac{J_R(\phi)^{1-\rho}}{1-\rho}\right)^{-1} z_R(\phi)$.

$\square$

Appendix 1.F   Robustness Checks

Table 1.5: Effect of past performance on the slope of the flow-performance relationship. Past performance is computed over 12 months. (Continues.)

| | F_{it+1} (Net Flow) | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{R}_{it}$	0.184***	0.177***	0.184***	0.177***	0.153***	0.147***
	(0.010)	(0.009)	(0.010)	(0.009)	(0.015)	(0.015)
$\text{PastPerf}_{i,[t-12,t-1]} \cdot \tilde{R}_{it}$	3.768***	3.285***	3.697***	3.175***		
	(1.141)	(1.186)	(1.029)	(1.023)		
$\tilde{R}_{it} \cdot \mathbb{I}[\text{PastPerf}_{i,[t-12,t-1]} > 0]$					0.062***	0.063***
					(0.019)	(0.020)
$\text{PastPerf}_{i,[t-12,t-1]} \cdot \mathbb{I}[\tilde{R}_{it} > 0]$					0.041	0.043
					(0.045)	(0.046)
$\text{PastPerf}_{i,[t-12,t-1]}$	0.432***	0.584***	0.432***	0.584***	0.293***	0.443***
	(0.035)	(0.039)	(0.035)	(0.039)	(0.047)	(0.051)
$(\text{PastPerf}_{i,[t-12,t-1]})^2$	0.427	0.660			−4.145	−0.528
	(2.017)	(2.246)			(9.092)	(9.521)
$(\text{CumPerf}_{i,[t-12,t]})^2$			0.502	0.774	6.694	2.852
			(2.368)	(2.636)	(11.395)	(12.105)
$\mathbb{I}[\tilde{R}_{it} > 0]$					−0.000	−0.000
					(0.000)	(0.000)
$\mathbb{I}[\text{PastPerf}_{i,[t-12,t-1]} > 0]$					0.002***	0.002***
					(0.000)	(0.000)
$\tilde{R}_{it}^2$	0.620***	0.687***	0.617***	0.682***	0.590***	0.694***
	(0.162)	(0.174)	(0.163)	(0.175)	(0.169)	(0.183)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Style-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Manager FE	No	Yes	No	Yes	No	Yes
Observations	153,127	153,127	153,127	153,127	153,127	153,127
R^2	0.351	0.383	0.351	0.383	0.351	0.383
Notes:	$^{*}p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$					

122

Table 1.5: (Continued.) The history dependence of the slope of the flow-performance relationship is measured by the coefficient on $\text{PastPerf}_{i[t-12,t-1]} \cdot \tilde{R}_{it}$. $\tilde{R}_{it}$ measures current performance and is calculated, for each month t, as the gross return of fund i in excess of the equally weighted average gross return of all funds with the same style. $\text{PastPerf}_{i[t-12,t-1]}$ measures the past performance of the manager and is calculated as the average excess return over the style benchmark in the twelve months from $t-12$ to $t-1$. The dependent variable, F_{it+1}, measures the net flow of capital and is calculated as the growth rate of assets under management from month t to month $t+1$ minus the net return over the same period. $\text{CumPerf}_{i[t-12,t]}$ is the average performance of the manager over the style benchmark in the months from $t-12$ to t. $\mathbb{I}[\cdot]$ is the indicator function. Controls include 12 lags of monthly net flows into the fund, the log of fund size, its expense ratio, the log of fund age, and the log of the manager's tenure. Standard errors are in parentheses and they are double-clustered at the fund and at the month level.
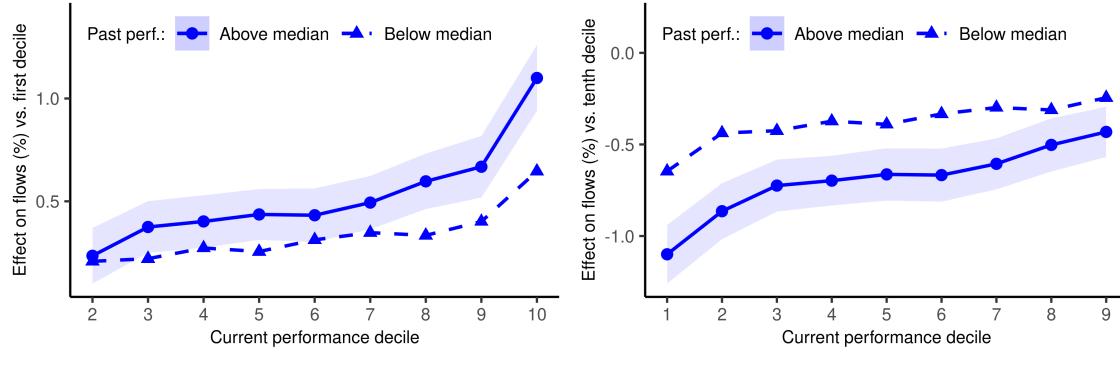
(a) Current good performance and flows (b) Current bad performance and flows

Figure 1.11: History dependence of the relation between flows and performance. Past performance is computed over 12 months. Figures (a) and (b) show how flows change with current performance and how the change depends on past performance. I sort funds into deciles based on their current performance and into halves based on their past performance. Past performance is the average excess return over the style benchmark in the previous 12 months. I then run regression of flows on dummies for the deciles of current performance, dummies for the halves of past performance, and interactions between the two sets of dummies. As controls, I include dummy variables for cumulative performance $\text{CumPerf}_{i[t-6,t]}$ sorted into deciles, 12 lagged flows, the logarithm of fund age, the logarithm of the manager's tenure, the logarithm of lagged assets under management, fund fees, fund-manager fixed effects, and style-month fixed effects. The shaded areas represent 95% confidence intervals for the change in the effect of current performance on flows when past performance increases above the median. Confidence intervals are constructed by double-clustering standard errors at the month and at the fund level.

In Figure (a), I plot the effect of current good performance (that is, performance relative to the first decile) on flows, while, in Figure (b), I plot the effect of current bad performance (that is, performance relative to the tenth decile) on flows.

(a) Effect of managerial tenure   (b) Effect of fund age

Figure 1.12: Effect of managerial tenure and fund age on the history dependence of the relation between flows and current performance. Past performance is computed over 12 months. (Continues.)

125

Figure 1.12: (Continued.) I run regression

$$
\begin{aligned}
F_{it+1} = \ & a_0 + a_1 \tilde{R}_{it} + a_2 \text{PastPerf}_{i[t-12,t-1]} \tilde{R}_{it} \\
& + a_3 \text{PastPerf}_{i[t-12,t-1]} + a_4 (\text{PastPerf}_{i[t-12,t]})^2 + a_5 \tilde{R}_{it}^2 \\
& \sum_{j=2}^{5} \text{TenureQuintile}_{it}^{j} \left(a_0^{Tj} + a_1^{Tj} \tilde{R}_{it} + a_2^{Tj} \text{PastPerf}_{i[t-12,t-1]} \tilde{R}_{it} \right. \\
& \left. \qquad + a_3^{Tj} \text{PastPerf}_{i[t-12,t-1]} + a_4^{Tj} (\text{PastPerf}_{i[t-12,t]})^2 + a_5^{Tj} \tilde{R}_{it}^2 \right) \\
& \sum_{j=2}^{5} \text{AgeQuintile}_{it}^{j} \left(a_0^{Aj} + a_1^{Aj} \tilde{R}_{it} + a_2^{Aj} \text{PastPerf}_{i[t-12,t-1]} \tilde{R}_{it} \right. \\
& \left. \qquad + a_3^{Aj} \text{PastPerf}_{i[t-12,t-1]} + a_4^{Aj} (\text{PastPerf}_{i[t-12,t]})^2 + a_5^{Aj} \tilde{R}_{it}^2 \right) \\
& \sum_{j=2}^{5} \text{SizeQuintile}_{it}^{j} \left(a_0^{Sj} + a_1^{Sj} \tilde{R}_{it} + a_2^{Sj} \text{PastPerf}_{i[t-12,t-1]} \tilde{R}_{it} \right. \\
& \left. \qquad + a_3^{Sj} \text{PastPerf}_{i[t-12,t-1]} + a_4^{Sj} (\text{PastPerf}_{i[t-12,t]})^2 + a_5^{Sj} \tilde{R}_{it}^2 \right) \\
& + c' X_{it} + \iota_{it}^{FMgr} + \iota_{it}^{SMon} + u_t,
\end{aligned}
$$

where $\text{TenureQuintile}_{it}^{j} = 1$ if, in month t, the tenure of the manager of fund i belongs to the jth quintile of the distribution of managerial tenure in month t; $\text{AgeQuintile}_{it}^{j} = 1$ if, in month t, the age of fund i belongs to the jth quintile of the distribution of fund age in month t; $\text{SizeQuintile}_{it}^{j} = 1$ if, in month t, the size of fund i belongs to the jth quintile of the distribution of fund size in month t. Table 1.3 contains the description of all the other variables used in the regression. Standard errors are double-clustered at the month and at the fund level.

In Figure (a), the dots in the figure represents estimated coefficients a_2^{Tj}'s. The vertical lines represent 90% confidence intervals for the incremental effect of tenure on the flow-performance slope relative to the first quintile: If the vertical red line at quintile j does not cross the dashed horizontal line, then we reject the hypothesis that $a_2^{Tj} \geq 0$ at a 95% confidence level.

Figure (b) is the analogous of Figure (a) for the effect of fund age on the history-dependence of the flow-performance relationship.

References

Agarwal, V., Daniel, N. D., and Naik, N. Y. (2009). Role of Managerial Incentives and Discretion in Hedge Fund Performance. *Journal of Finance*, 64(5):2221–2256.

Almazan, A., Brown, K. C., Carlson, M., and Chapman, D. A. (2004). Why constrain your mutual fund manager? *Journal of Financial Economics*, 73(2):289–321.

Alvarez, O., Lasry, J.-M., and Lions, P.-L. (1997). Convex viscosity solutions and state constraints. *Journal de Mathématiques Pures et Appliquées*, 76(3):265–288.

Basak, S. and Pavlova, A. (2013). Asset Prices and Institutional Investors. *American Economic Review*, 103(5):1728–1758.

Bergemann, D. and Hege, U. (2005). The Financing of Innovation: Learning and Stopping. *RAND Journal of Economics*, 36(4):719–752.

Berk, J. B. and van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1):1–20.

Berk, J. B. B. and Green, R. C. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, 112(6):1269–1295.

Bhattacharya, S. and Pfleiderer, P. (1985). Delegated portfolio management. *Journal of Economic Theory*, 36(1):1–25.

Biais, B., Mariotti, T., Rochet, J.-C., and Villeneuve, S. (2010). Large Risks, Limited Liability, and Dynamic Moral Hazard. *Econometrica*, 78(1):73–118.

BIS (2003). Incentive structures in institutional asset management and their implications for financial markets. Technical report, Bank For International Settlements.

Brown, K. C., Harlow, W. V., and Starks, L. T. (1996). Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry. *Journal of Finance*, 51(1):85–110.

Cadenillas, A., Cvitanić, J., and Zapatero, F. (2007). Optimal risk-sharing with effort and project choice. *Journal of Economic Theory*, 133(1):403–440.

Chen, Q., Goldstein, I., and Jiang, W. (2010). Payoff complementarities and financial fragility: Evidence from mutual fund outflows. *Journal of Financial Economics*, 97:239–262.

Chen, Q. I., Goldstein, I., and Jiang, W. (2008). Directors' Ownership in the U.S. Mutual Fund Industry. *Journal of Finance*, 63(6):2629–2677.

Chevalier, J. and Ellison, G. (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 105(6):1167–1200.

Chevalier, J. and Ellison, G. (1999). Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance. *Journal of Finance*, 54(3):875–899.

Cisternas, G. (2018). Two-Sided Learning and the Ratchet Principle. *Review of Economic Studies*, 85(1):307–351.

Cuoco, D. and Kaniel, R. (2011). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics*, 101(2):264–296.

128

Del Guercio, D. and Tkac, P. A. (2002). The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds vs. Pension Funds. *Journal of Financial and Quantitative Analysis*, 37(4):523–557.

DeMarzo, P. M. and Fishman, M. J. (2007). Agency and Optimal Investment Dynamics. *Review of Financial Studies*, 20(1):151–188.

DeMarzo, P. M., Fishman, M. J., He, Z., and Wang, N. (2012). Dynamic Agency and the q Theory of Investment. *Journal of Finance*, 67(6):2295–2340.

DeMarzo, P. M. and Sannikov, Y. (2006). Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model. *Journal of Finance*, 61(6):2681–2724.

DeMarzo, P. M. and Sannikov, Y. (2017). Learning, Termination, and Payout Policy in Dynamic Incentive Contracts. *Review of Economic Studies*, 84:182–236.

Di Tella, S. (2017). Optimal Regulation of Financial Intermediaries. Technical report.

Di Tella, S. and Sannikov, Y. (2018). Optimal Asset Management Contracts with Hidden Savings.

Dybvig, P. H., Farnsworth, H. K., and Carpenter, J. N. (2010). Portfolio Performance and Agency. *Review of Financial Studies*, 23(1):1–23.

Elton, E. J., Gruber, M. J., and Blake, C. R. (2001). A First Look at the Accuracy of the CRSP Mutual Fund Database and a Comparison of the CRSP and Morningstar Mutual Fund Databases. *Journal of Finance*, 56(6):2415–2430.

Evans, R. B. (2010). Mutual Fund Incubation. *Journal of Finance*, 65(4):1581–1611.

Fama, E. (1980). Agency problems and the theory of the firm. *Journal of Political Economy*, 88(2):270–282.

Fama, E. F. and French, K. R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *Journal of Finance*, 65(5):1915–1947.

Franzoni, F. and Schmalz, M. C. (2017). Fund Flows and Market States. *Review of Financial Studies*, 30(8):2621–2673.

Getmansky, M., Lo, A. W., and Makarov, I. (2004). An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529–609.

Gibbons, R. and Murphy, K. J. (1992). Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy*, 100(3):468–505.

Goldstein, I., Jiang, H., and Ng, D. T. (2017). Investor flows and fragility in corporate bond funds. *Journal of Financial Economics*, 126(3):592–613.

Halac, M., Kartik, N., and Liu, Q. (2016). Optimal Contracts for Experimentation. *Review of Economic Studies*, 83(3):1040–1091.

Han, B., Hirshleifer, D., and Walden, J. (2018). Social Transmission Bias and Investor Behavior.

He, Z. (2012). Dynamic Compensation Contracts with Private Savings. *Review of Financial Studies*, 25(5):1494–1549.

He, Z., Wei, B., Yu, J., and Gao, F. (2017). Optimal Long-Term Contracting with Learning. *Review of Financial Studies*, 30(6):2006–2065.

He, Z. and Xiong, W. (2013). Delegated asset management, investment mandates, and capital immobility. *Journal of Financial Economics*, 107(2):239–258.

Heinkel, R. and Stoughton, N. M. (1994). The Dynamics of Portfolio Management Contracts. *Review of Financial Studies*, 7(2):351–387.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66(1):169–182.

Hörner, J. and Samuelson, L. (2013). Incentives for experimenting agents. *RAND Journal of Economics*, 44(4):632–663.

Huang, J., Wei, K. D., and Yan, H. (2007). Participation Costs and the Sensitivity of Fund Flows to Past Performance. *Journal of Finance*, 62(3):1273–1311.

Hunter, D., Kandel, E., Kandel, S., and Wermers, R. (2014). Mutual fund performance evaluation with active peer benchmarks. *Journal of Financial Economics*, 112(1):1–29.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., and Vestman, R. (2018). Are Mutual Fund Managers Paid for Investment Skill? *Review of Financial Studies*, 31(2):715–772.

Ippolito, R. A. (1992). Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry. *Journal of Law and Economics*, 35.

Kaniel, R. and Kondor, P. (2013). The Delegated Lucas Tree. *Review of Financial Studies*, 26(4):929–984.

Kaplan, S. N. and Schoar, A. (2005). Private Equity Performance: Returns, Persistence, and Capital Flows. *Journal of Finance*, 60(4):1791–1823.

Katsoulakis, M. A. (1994). Viscosity Solutions of Second Order Fully Nonlinear Elliptic Equations with State Constraints. *Indiana University Mathematics Journal*, 43(2):493–519.

Khorana, A., Servaes, H., and Wedge, L. (2007). Portfolio manager ownership and fund performance. *Journal of Financial Economics*, 85(1):179–204.

Laffont, J.-J. and Tirole, J. (1990). Adverse Selection and Renegotiation in Procurement. *Source: The Review of Economic Studies*, 57(4):597–625.

Lynch, A. W. and Musto, D. K. (2003). How Investors Interpret Past Fund Returns. *Journal of Finance*, 58(5):2033–2058.

Ma, L., Tang, Y., and Gomez, J. (2019). Portfolio Manager Compensation in the U.S. Mutual Fund Industry. *Journal of Finance*, 74(2):587–638.

Miao, J. and Zhang, Y. (2015). A duality approach to continuous-time contracting problems with limited commitment. *Journal of Economic Theory*, 159:929–988.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.

Ou-Yang, H. (2003). Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem. *Review of Financial Studies*, 16(1):173–208.

Palomino, F. and Prat, A. (2003). Risk Taking and Optimal Contracts for Money Managers. *RAND Journal of Economics*, 34(1):113–137.

Panageas, S. and Westerfield, M. M. (2009). High-Water Marks: High Risk Appetites? Convex Compensation, Long Horizons, and Portfolio Choice. *Journal of Finance*, 64(1):1–36.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, 116(1):23–45.

Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*. Springer-Verlag.

Prat, J. and Jovanovic, B. (2014). Dynamic contracts when the agent's quality is unknown. *Theoretical Economics*, 9(3):865–914.

Sannikov, Y. (2008). A Continuous-Time Version of the Principal–Agent Problem. *Review of Economic Studies*, 75(3):957–984.

Sannikov, Y. (2014). Moral Hazard and Long-Run Incentives.

Sirri, E. R. and Tufano, P. (1998). Costly Search and Mutual Fund Flows. *Journal of Finance*, 53(5):1589–1622.

Spiegel, M. and Zhang, H. (2013). Mutual fund risk and market share-adjusted fund flows. *Journal of Financial Economics*, 108(2):506–528.

Strulovici, B. (2011). Renegotiation-Proof Contracts with Moral Hazard and Persistent Private Information.

Vayanos, D. and Woolley, P. (2013). An Institutional Theory of Momentum and Reversal. *Review of Financial Studies*, 26(5):1087–1145.

Williams, N. (2009). On Dynamic Principal-Agent Problems in Continuous Time.

Williams, N. (2011). Persistent Private Information. *Econometrica*, 79(4):1233–1275.

Zheng, L. (1999). Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability. *Journal of Finance*, 54(3):901–933.

CHAPTER 2

THE TRANSMISSION OF QUANTITATIVE EASING TO CORPORATE BOND PRICES AND ISSUANCE

This chapter is a joint work with Mattia Montagna (DG MF Systemic Risk and Financial Institutions, European Central Bank, Frankfurt am Main, Germany). The views expressed in this chapter are those of the authors and do not necessarily reflect those of the European Central Bank.

2.1 Introduction

Since the financial crisis of 2008, the traditional tools of monetary policy have been challenged by market segmentation, financial instability, and low interest rates. As a consequence, central banks in advanced economies have become increasingly reliant on unconventional measures, among which quantitative easing (QE) features a primary role. However, despite the large scale of such programs, surprisingly little research has investigated their effects on the financing activities of firms and the channels through which they affect the real sector of the economy. By providing extensive and robust evidence of the transmission channels of QE, this paper contributes to filling the gap between monetary policy and economic theory and highlights the key empirical facts that theoretical papers in this area should try to match.

We exploit the European Central Bank (ECB)'s Corporate Sector Purchase Program (CSPP) to study the transmission channels of QE on the cross section of corporate bond prices and firms' financing activity. A number of papers have analyzed

135

the effects of asset purchases by looking at the yields of the securities that past QE programs targeted. However, previous contributions did not compare the targeted securities to reasonable control groups, nor did they investigate the effects of QE on security issuance. Therefore, such studies provided only suggestive evidence on the transmission channels of QE, because they lacked counter-factual observations and omitted any information contained in bond issuance decisions. Our approach exploits the cross-sectional heterogeneity of bonds and firms in the euro area and allows us to disentangle the marginal contribution of each transmission channel while controlling for a set of bond- and firm-level characteristics. We therefore identify channels of transmission on the basis of appropriate counterfactual evidence. Moreover, by looking at the effect of QE on issuance, we are able to establish the presence of transmission channels that we could not have observed in bond prices alone.

In our identification strategy, we rely on the particular rules that govern the eligibility of a bond for the CSPP. Through the CSPP, the ECB can purchase euro-denominated bonds issued by non-bank corporations domiciled in the euro area. On the day of the announcement, the ECB identified a set of securities that are eligible for purchase and a set of non-eligible securities. More specifically, the ECB was allowed to buy only securities that were accepted as collateral for its refinancing operations. This policy rule gives us a chance to explore the heterogeneous impact of QE on different securities based on their eligibility.

Besides the rules that govern eligibility, the past price performance of eligible and non-eligible bonds also strengthens our identification strategy. Before the announcement of the program, the market for *eligible* corporate bonds in the euro area was

not showing any sign of distress. Instead, in the months leading up to the CSPP announcement, prices declined substantially for non-eligible bonds. Thanks to the possibility of clearly separating the effect on eligible securities and the effect on distressed securities, we are able to disentangle whether QE mostly affected securities on the basis of their eligibility or their distress.

Our results challenge conventional wisdom on the transmission channels of unconventional monetary policy.[1] In particular, we find that after the announcement, the price of eligible securities increased *less* than the price of non-eligible securities, that the issuance of eligible securities increased more than the issuance of non-eligible securities, and that eligible issuers (that is, firms that issue eligible bonds) did not increase issuance more than non-eligible issuers. Whereas heterogeneity in price change is mainly observed across different firms, rather than across different bonds within the same firm, heterogeneity in bond issuance is observed within firms that substitute eligible issuance for non-eligible issuance.

The first transmission channel we consider, the scarcity channel, implies that QE corresponds to an increase in the aggregate demand for eligible bonds. Previous studies, which ignored the possibility of issuers supplying more eligible bonds, then argued that the ultimate outcome should be an increase in the price of eligible bonds. Our results demonstrates the limited power of such studies. We find that a scarcity channel is present and that it manifests itself through eligible firms increasing the issuance of eligible bonds compared to non-eligible bonds. In fact, firms are natural

---

1. Bernanke et al. (2012) explain the leading views of central bankers on the transmission channels of monetary policy and stress the role of asset scarcity and portfolio rebalancing as key mechanisms in explaining the effect of QE on relative asset prices.

137

short sellers of their own securities so that, if issuance frictions are limited, the scarcity channel should only marginally affect relative prices, and should instead primarily affect relative bond quantities. In equilibrium, firms should issue bonds of any certain class up to the point where investors' demand is satiated and no relative price premium is present across different bond classes.

We document that differences in price changes of bonds after the CSPP announcement are mainly driven by differences in their previous price performance, thus suggesting that risk is the primary channel of transmission to prices. Before the CSPP announcement, several firms experienced large drops in their debt valuation. After the announcement, these firms were the ones that enjoyed the largest price increases. However, these firms did not increase issuance in the months after the announcement. Although seemingly contradictory, these two findings can be reconciled in a framework where firms may take inefficiently high levels of risk.[2] If QE is able to reduce the incentives for inefficient risk-taking, we might simultaneously observe an increase in debt prices and a negligible, or even negative, change in issuance. As another possible explanation, we might expect these riskier firms to be facing a binding borrowing constraint or to be operating below capacity. Although we control for credit ratings in our empirical analysis, we chose to assess a firm's riskiness on the basis of its past price performance because ratings may respond with a significant lag to negative news and because prices may reflect risks and sources of uncertainty that are not perfectly captured by credit ratings.[3]

---

2. Stein (2012), Greenwood et al. (2016), and Woodford (2016) provide models of excessive risk taking.

3. Rating agencies' downgrades are known to lag the market (Hull et al., 2004; Tang, 2009).

If a liquidity channel is present, then, when the central bank increases the supply of liquid assets by purchasing bonds with reserves, we should expect the prices of more liquid assets to decline. We exclude the presence of such a channel on the basis of two pieces of evidence. First, although eligible bonds are more liquid that non-eligible bonds because, if we control for past price performance the effect of QE on the price of eligible bonds is either negligible or even positive. Second, we repeat all our analyses using the announcement of the Public Sector Purchase Program (PSPP) of the ECB. This program involved a much larger supply of liquidity, and yet we are unable to find any statistically significant difference across the price changes for eligible and non-eligible corporate bonds.

We conduct our analyses using a novel data source that contains information on all securities issued in the euro area and that is compiled and managed by the Eurosystem. We focus on euro-denominated debt securities issued by non-financial corporations (NFCs). On the one hand, we restrict the sample to euro-denominated bonds because we would like to avoid concerns related to exchange-rate movements. On the other hand, we exclude financial institutions because QE directly affects their investment opportunities and we would therefore struggle to identify channels of transmission that operate through changes in the cost of borrowing, which is the main objective of this paper.

We are especially interested in how QE changes firms' cost of borrowing, because the ECB itself identified it as a key channel to transmit monetary-policy actions to the real economy. For example, in the January 22, 2016, the ECB announced the PSPP in a press release stating that

"Asset purchases provide monetary stimulus to the economy in a context where key ECB interest rates are at their lower bound. They further ease monetary and financial conditions, making access to finance cheaper for firms and households. This tends to support investment and consumption, and ultimately contributes to a return of inflation rates towards 2%".

The remainder of the paper is organized as follows. In section 2.2, we discuss the transmission channels of QE identified by the theoretical literature. In section 2.3, we relate our paper to previous empirical literature. In section 2.4, we provide background information on the euro area corporate bond market and the institutional details of the CSPP. In section 2.5, we describe the data used in this paper. In section 2.6, we explain the details of our econometric analyses by presenting the variables of interest and the empirical strategies. In section 2.7, we assess the implications of our cross-sectional analyses as tests for the presence of scarcity, liquidity, and risk channels. We conclude the paper with section 2.8.

Appendices contain further information and all plots and tables. Appendix 2.A contains further details on all the QE programs of the ECB. Appendix 2.B contains all plots mentioned in the main body of the paper and related to the CSPP announcement. Appendix 2.C contains all tables used for the analysis of the price impact of the CSPP. Appendix 2.D contains all tables used for the analysis of the issuance around the CSPP announcement. Finally, Appendix 2.E contains analogous plots and tables for the announcement of the PSPP.

2.2 Theoretical Framework

To better explain the implications of our empirical results, we first discuss the main transmission channels of unconventional monetary policy that the theoretical literature has suggested. To facilitate the explanation, consider the basic one-period asset-pricing equation,

$$P_{it} = \mathrm{E}[\pi_{t,t+1} D_{i,t+1} | \mathcal{F}_t], \tag{2.1}$$

where P_{it} is the price of security i at time t, $\pi_{t,t+1}$ is the stochastic discount factor (SDF) from time $t + 1$ to t, $D_{i,t+1}$ is the payoff of the security at time $t + 1$, and $\mathcal{F}_t$ is the time-t σ-algebra capturing the available information. By simply looking at this equation, we can clearly identify three possible ways in which monetary policy can affect asset prices: (i) by changing the distribution of $\pi_{t,t+1}$, (ii) by changing the distribution of D_{t+1}, or (iii) by changing investors' information $\mathcal{F}_t$.

How does this simple equation help us understand the effects of QE? Consider, for example, a frictionless consumption-based asset-pricing model, such as the one in Wallace (1981). In this model, QE is financed by non-distortionary lump-sum taxes, which thus leave $\pi_{t,t+1}$ unaltered. Moreover, asset purchases do no directly affect the distribution of the asset's payoff nor the information available to investors. The conclusion that, in this setting, QE is neutral is therefore not surprising.

To overcome the neutrality of QE that is embedded in any frictionless model, economists have developed a series of models in which QE does have an effect on asset prices and, consequently, on the allocation of credit and resources. In these models, unconventional monetary policy activates one or more channels of transmis-

sion, leading to a change either in the relative price of financial instruments or in the wedge between market prices and fundamental values, or both.

Whereas most models analyze the impact of QE on asset prices while keeping supply fixed, in the real world, firms are natural short sellers of their own securities. Therefore, if they act as neoclassical arbitrageurs, the presence of some transmission channels may simply result in a change in the firms' liability composition rather than in the relative prices of financial instruments. This possibility must be kept in mind in order to fully understand the consequences of QE and interpret them within the framework of economic theory.

The channels of transmission identified by the theoretical literature are the scarcity channel, the liquidity channel, the risk channel, and the signaling channel. We discuss them below.

Scarcity channel

A scarcity channel arises whenever securities are valued for reasons beyond the flow of consumption that they provide. In this case, the SDF $\pi_{t,t+1}$ depends not only on the consumption stream of the investors, but also on their portfolio composition. As a central bank purchases assets, those assets become scarcer and their valuation increases. This channel therefore implies that, everything else kept constant, the prices of securities purchased by the central bank should increase. Moreover, whereas a scarcity channel should manifest itself only when the central bank starts purchasing assets and thus making them scarcer, the presence of (limited) arbitrage possibilities for some agents may imply that most of the price impact will be observed at the

announcement.

Models of this type date back to Tobin (1969), who assumes that investors have preferences over assets, as if assets were consumption goods. More recently, new models have been developed that feature agents with a "preferred habitat," that is, agents who have preferences for assets with some particular maturity or risk profile, as in Vayanos and Vila (2009), Krishnamurthy and Vissing-Jorgensen (2012), and Greenwood et al. (2010). Although these models offer many degrees of freedom in the choice of the assets investors exogenously value and of the extent to which investors value such assets, some of them endogenously generate preferences for assets. For example, in Lenel (2017), assets are valued depending on the extent to which they can be used as collateral for borrowing.

We can can thus summarize an empirical prediction of the scarcity channel as follows.

Empirical prediction (1): A scarcity channel implies that, following QE, the price of eligible securities should increase relative to other securities, *if supply is fixed.*

However, this prediction may not hold if some investors do not have preferences over assets and have short-selling opportunities. In this case, prices will always reflect the fundamental value of the underlying cash flow, whereas the quantity of assets supplied to habitat investors will adjust to align the assets' prices to the valuation of the short sellers. Because firms are natural short sellers of their own debt, if they are unconstrained, the presence of investors with preferences for assets implies that

quantities, rather than prices, will adjust in response to QE. We can then obtain a second empirical prediction of the scarcity channel.

Empirical prediction (2): If firms are unconstrained in the capital market, the quantity of eligible securities will increase relative to non-eligible securities, leaving relative prices unchanged.

Liquidity channel

A liquidity channel is a specific form of scarcity channel and it takes into consideration the fact that QE involves the central bank swapping securities for reserve, thus increasing the supply of highly liquid assets. If investors value assets for their liquidity, QE will then lower liquidity premia, thus causing the price of liquid assets to decline relative to the price of less liquid assets. These effects have been captured by models as in Drechsler et al. (2014) and Bianchi and Bigio (2014), where investors face liquidity shocks and need to keep a buffer of liquid assets to avoid costly liquidation.

The empirical prediction of the liquidity channel is therefore the following.

Empirical prediction: A liquidity channel implies that, following QE, the price of more liquid securities decreases relative to the price of less liquid securities.

Risk channel

QE may affect asset prices also by changing investors' valuation of risk or the amount of economic risk itself. For example, in many intermediary asset-pricing models, QE is non-neutral if the central bank, by swapping risky assets for riskless assets, relaxes the balance-sheet constraints of intermediaries and modifies their consumption stream. In this case, a risk channel works mainly through changes in risk premia, driven by changes in the distribution of the SDF $\pi_{t,t+1}$. Several papers in this area, such as He and Krishnamurthy (2013), Brunnermeier and Sannikov (2016), Curdia and Woodford (2011), Gertler et al. (2010), and Gertler and Karadi (2011), view QE essentially as a distortionary taxation that modifies the consumption stream of marginal investors, thus changing equilibrium risk prices.

Whereas most theoretical contributions focus on the effect on monetary policy on risk prices, we can also entertain the possibility that monetary policy may directly affect the quantity of risk in the economy, and hence the distribution of the payoff D_{t+1}. This possibility is reasonable if, as argued in Stein (2012), Greenwood et al. (2016), and Woodford (2016), central bank interventions directly address market failures that would lead to excessive risk-taking.

Therefore, a first empirical prediction of the risk channel relates QE to changes in prices of risky securities.

Empirical prediction (1): A risk channel predicts that, following QE, the price of riskier securities increases relative to the price of less risky security, all else being equal.

Moreover, whenever risk prices decrease, the cost of capital of risky firms decreases as well, thus stimulating more issuance. This consideration leads to a second prediction.

Empirical prediction (2): If a risk channel acts through risk prices, then, following QE, the issuance of riskier firms should increase relative to the issuance of less risky firms, all else being equal.

Signaling channel

According to the last transmission channel that we discuss, namely, the signaling channel, unconventional monetary policy conveys new information to investors. Specifically, monetary authorities may possess superior information on financial markets and the real economy, and they might reveal such information when they announce new policy measures. However, this informational effect may either increase or decrease asset prices, because investors could interpret the policy either as a negative signal that the economy is performing badly, or as a positive signal that some securities are considered of high enough quality for a central bank to purchase them. Furthermore, unconventional monetary policy may also signal the willingness of authorities to step in to prevent major disruptions in financial markets, thus reducing policy uncertainty. Finally, as in Eggertsson et al. (2003), Clouse et al. (2003), and Bhattarai et al. (2015), asset purchases may serve as a credible commitment for the central bank to keep low interest rates in the future when the short-term rate

has already reached its lower bound. This last signaling effect implies that QE announcement should flatten the term structure. Although we do not directly test for the presence of a signaling channel, we do control for movements of term structure of interest rates.

2.3   Related Empirical Literature

Several papers have analyzed the impact of QE on bond prices and yields. However, previous QE events lacked one of the key features of the CSPP, namely, the ability to clearly distinguish distressed securities from eligible securities. Therefore, many of the previous contributions could not clearly disentangle the effects of a scarcity channel and a risk channel. Researchers often claim to find evidence of a scarcity channel, but also acknowledge the peculiar market circumstances in which they found it.

Krishnamurthy and Vissing-Jorgensen (2011) and Krishnamurthy and Vissing-Jorgensen (2013) find that the QE announcements by the Fed mainly affected the yields of the securities that the program targeted, with limited effects on other securities. This finding is consistent with the presence of a scarcity channel. However, as they highlight, the Fed launched a mortgage-backed security (MBS) purchase program in a period when the MBS market was particularly distressed and prices had previously dropped.

Similar to our paper, D'Amico et al. (2012) and D'Amico and King (2013) explore the cross-sectional impact of QE announcements using security-level data. They focus on government bonds and document the heterogeneity in the response of different

securities depending on their maturity and whether they are eligible for the program. They find evidence of a scarcity channel, which is present even at the time of the purchase, even though the effects are temporary. However, they also note that their findings may be, at least in part, the result of purchases taking place in a period of particular financial distress and low market liquidity. Gagnon et al. (2011) find that QE announcements lowered the yields of long-term securities but attribute this change in yields mostly to a reduction in risk premia, rather than a reduction in the expected path of short-term rates.

Similar to our study, Abidi et al. (2017) focus on the CSPP program and document that it had the largest price effect on non-eligible bonds. They therefore conclude that non-eligible bonds benefited the most from the program. Instead, we highlight debt value increased the most for distressed bonds. Indeed, as we disentangle the marginal contribution of each transmission channel of QE and employ a more comprehensive dataset, we are able to separately identify the role of eligibility and financial distress.

Hamilton and Wu (2012) and Greenwood and Vayanos (2014) adopt the model of Vayanos and Vila (2009) to quantify the differential impact of QE on the prices of assets of different maturities. Both contributions find that the relative supply of bonds of different maturities is associated with relative prices and expected returns. Hamilton and Wu (2012) stress, in particular, how effects tend to be stronger when the interest rate is at the lower bound, whereas Greenwood and Vayanos (2014) highlight how supply effects are more pronounced when risk aversion is higher. Both elements (very low rates and high risk aversion), however, are likely to coexist in

periods of financial distress. This observation highlights that the non-neutrality of QE on prices may be tightly linked to markets being impaired.

Many other papers have looked at the QE announcements by the Fed, the ECB, the Bank of England, and the Bank of Japan (Altavilla et al., 2015; Andrade et al., 2016; Falagiarda and Reitz, 2015; Fratzscher et al., 2016; Joyce et al., 2011; Lam, 2011; Swanson, 2011, 2015; Szczerbowicz et al., 2015; Ueda, 2012), and find that asset purchases have a significant effect on yields, that most of this effect is observed at the announcement, that eligible assets experience the greatest price increases, and that long-maturity bonds are more significantly affected by QE. None of these studies however, systematically employed counterfactual analysis.

By contrast, our cross-sectional analysis establishes that scarcity does not play a role in the price reaction to QE. Rather, market impairment is the leading condition for QE to exert a positive impact on security prices.

We also move beyond this stream of literature and look at the issuance of bonds by NFCs in the euro area, and find convincing evidence that the price impact of a scarcity channel is neutralized by firms substituting across sources of financing. Once again, CSPP provides a unique identification opportunity. Indeed, contrary to the Fed program, CSPP targets private entities that are not government entities or government-sponsored enterprises and are thus more likely to respond to changes in the external environment in a profit-maximizing way. Moreover, we can once again rely on the large cross section of European firms to develop counterfactual analysis.

In the QE literature, Di Maggio et al. (2016) conduct an analysis on quantities that is similar to ours. They study mortgage origination after the first round of QE

in the US, and find that mortgages that were eligible for purchase by the Fed were originated in larger quantities than non-eligible mortgages.

An established empirical literature tries to identify the effects of capital market conditions on the determination of firms' liability structure (Baker et al., 2003; Baker and Wurgler, 2000; Faulkender, 2005; Faulkender and Petersen, 2006). This literature consistently finds that an increase in firms' issuance of liabilities of a certain class is able to predict low returns for the same instrument in the future, suggesting that firms time the market in order to take advantage of the high valuation of certain securities.

To the extent that QE introduces a demand shock in the capital market that lower yields of certain asset classes, our paper can bring insights into this area of the academic debate. Indeed, QE can be interpreted as a quasi-experimental demand shock. The changes that we observe in the firms' issuance policy hint toward an attempt by managers to actually time the market.

In this literature, whether managers are actually successful at reducing their cost of capital is still debated (Butler et al., 2006). Although we do not fully tackle this issue, because doing so would require us study the effects of QE on bank loans and equity, we provide suggestive evidence that QE does not reduce the cost of borrowing. Indeed, eligible firms do not increase their total issuance compared to non-eligible firms, even though they do substitute across forms of bond financing. This behavior is symptomatic of an equilibrium where, regardless of the preferences investors have over assets, firms act as arbitrageurs that prevent market valuation from deviating too much from fundamentals.

2.4 Background Information on the Euro Area Corporate Bond Market and the CSPP

Starting in October 2014, the ECB launched a series of four asset-purchase programs (APP) that, combined, contributed to a total of €1,896bn of the ECB portfolio as of May 2017. Such programs are the third Covered Bond Purchase Program (CBPP3), the Asset-Back Securities Purchase Program (ABSPP), the Public Sector Purchase Program (PSPP) and the Corporate Sector Purchase Program (CSPP).

In this section, we highlight the key features of the CSPP, which is the focus of our paper. We briefly describe the other programs in Appendix 2.A, which also contains details on the various changes that have been made to the programs' duration and size.

The CSPP targets euro-denominated bonds issued by euro-area private corporations, excluding credit institutions. Because we are interested in NFCs, Figure 2.1 plots the outstanding amount of euro-denominated bonds issued by NFCs in the 19 countries of the euro area. Toward the end of the period, the size of the market is €958bn,[4] which is approximately 10% of the nominal GDP in 2016, which amounts to €10,745bn.[5] Europe has a mainly bank-based financial system and its corporate bond market is not particularly big relative to that in the US, where NFCs' bonds amount is approximately 30% of the nominal GDP.[6]

---

4. Source: Centralized Security Database. See section 2.5.

5. Source: ECB's Statistical Data Warehouse, http://sdw.ecb.europa.eu/.

6. At the end of 2016, the outstanding amount of NFC bonds in the USA was $5,075bn and the nominal GDP in the same year was $18,569bn. Source: St. Louis Fed's FRED, https://fred.stlouisfed.org/.

On top of this general background, the CSPP was announced in a period that was characterized by high levels of distress in the corporate bond market. Because the we study the effects of the CSPP announcement on distressed bonds, we discuss the situation of the euroa area corporate bond market before presenting the institutional details of the CSPP.

2.4.1   The Corporate Bond Market before the CSPP Announcement

The corporate bond market in Europe was going through a period of price declines in the two months leading up to the CSPP announcement. In January 2016, The Wall Street Journal wrote, "A wave of selling has taken Europe's corporate-bond market to levels typically seen during recessions, another indication that the turmoil in global markets could spread into the wider economy" (Whittall, y 18). Other newspaper articles can later be found discussing the low valuation of corporate bonds at the beginning of 2016 (Barley, ry 3; Platt, y 12; Smith, ry 5). Signs of distress in the corporate bond market are also discussed in the February and March 2016 Economic Bulletins of the ECB (ECB, 2016a,b).

Figure 2.2 in Appendix 2.B plots the average daily log prices and log returns for bonds with ratings between BBB+ and BB, and Figure 2.3 provides separate series for eligible and non-eligible bonds in this rating range. From these figures, non-eligible bonds appear to have experienced the largest price increase following the announcement. However, non-eligible bonds also appear to have been going thorough a period of intense pressure in the months before the CSPP announcement.

To gain further insight, we consider the performance of bonds before the an-

nouncement of the CSPP. For each bond, we consider the log-price change from December 10, 2016, to February 11, 2016 (four weeks before the announcement). We then group bonds into deciles according to their log-price change in this period. We then identify a group of distressed bonds (bonds in the first decile), a group of mildly distressed bonds (bonds in the second decile), and a group of non-distressed bonds (bonds in the third decile or higher).

Figure 2.4 show the behavior of bond prices and returns when they are grouped by level of distress. The only series that experiences a substantial jump around the CSPP announcement are the ones of the bonds in a state of distress.

These patterns cannot be observed around other QE announcements. Indeed, in Appendix 2.E, we repeat the same exercises for the PSPP announcement. In this case, eligible and non-eligible bonds do not seem to display any heterogeneity in their response to this monetary-policy announcement. Rather, their prices move in parallel fashion, which suggests that liquidity premia should not play a role, because the PSPP involved a much larger supply of reserves. However in the month preceding the PSPP announcement, no significant decline occurred in bond prices.

To our knowledge, no paper rigorously attempts to identify the causes of the corporate bond price decline at the beginning of 2016, and investigating them is well beyond the scope of our paper. However, it suffices to highlight that the European corporate bond market was, at that time, going through a period of high downward pressure and increasing uncertainty.

Although the ECB policy decision may well have been an endogenous response to this situation, this possibility does not detract from our analysis. On the contrary,

153

we can exploit the institutional details of the CSPP as well as the price history of different bonds in the months leading up to the announcement to better identify the channels through which QE works.

2.4.2   Institutional Details on the CSPP

The peculiarity of the CSPP, compared to all the QE programs undertaken by the Fed, is that it directly targets corporate securities, as opposed to bonds issued by governments or government-sponsored entities. This peculiarity provides an ideal framework to test the transmission channels of monetary policy to the real economy, because it establishes a direct link between central bank purchases and real economic entities. Indeed, the stated aim of the program is to reinforce the pass-through of monetary-policy stimuli to the real economy by directly acquiring marketable debt of private corporations.

The program was announced on March 10, 2016, and purchases started on June 8, 2016. At the time of the announcement, the ECB was already conducting the CBPP3, ABSPP, and PSPP, totaling €60bn each month combined, with the PSPP constituting the great majority of the total. On March 10, 2016, the ECB announced the intention to introduce the CSPP and raise the combined amount of bond purchases to €80bn each month starting from April 2016, whereas CSPP purchases were said to begin toward the end of the second quarter of the same year. Purchases were intended to last until March 2017, or beyond if necessary.

Contrary to the Fed, the ECB has always accepted corporate bonds as collateral for its refinancing operations, provided that they satisfy a list of eligibility criteria.

These criteria must be satisfied also in order for a bond to be eligible for CSPP purchases, together with some additional requirements.

Specifically, in order to be eligible as collateral, securities must satisfy the following requirements: (1) They must be debt instruments, (2) they must be euro-denominated and (3) issued in the euro area, (4) the issuer must be established in the euro area or in a G10 country, and (5) the securities must have a rating of at least BBB- (or equivalent), for long-term debt, or A-2 (or equivalent) for short-term debt. A list of bonds that are eligible as collateral is published daily on the ECB website.

In addition to these requirements, the ECB set further restrictions on the bonds that can be purchased under the CSPP program. In particular, bonds issued by credit institutions and by entities whose ultimate parent is a bank are excluded from the set of eligible securities. Moreover, the issuer must be domiciled in the euro area.

On April 21, 2016, the ECB released some additional technical details that, among other things, specify that bonds, in order to be eligible for CSPP, must have a maturity between 6 months and 31 years. The ECB made this choice in order to avoid too frequent portfolio adjustments, while also including in the set of eligible securities the bonds of smaller firms that finance themselves with short-term liabilities.

The eligibility criteria must be satisfied at the time of the purchase. If a bond loses eligibility status (e.g., because of a downgrade), the ECB is not required to sell it. Moreover, as for all the other asset-purchase programs, the ECB will reinvest all the capital repayments until reinvestment is deemed appropriate.

The size of the program is not negligible. Indeed, as of May 2017, the ECB

155

held a portfolio of €89.8bn bonds purchased under the CSPP, with average monthly purchases of €7.5bn.[7] These numbers compare with the €887bn total outstanding amount of bonds issued by NFCs in the euro area as of February 2016, of which €500bn were eligible as collateral at that time.

Purchases are carried out both in the primary and the secondary market, with the exception of bonds issued by public undertakings that are purchased exclusively in the secondary market. As of May 2017, 85.34% of the CSPP holdings have been purchased on the secondary market.

The ECB regularly discloses the monthly purchases and holding for each program. It also publishes the list of the securities that are held in its portfolio and makes such securities available for security lending. The amount of individual securities purchased is, however, not disclosed.

2.5 Data

The main source of data is the Centralized Security Database (CSDB). The CSDB contains security-level information on every equity, debt, and hybrid instrument issued by euro-area residents. This dataset is managed by the Eurosystem and is updated at a monthly frequency with observations starting in February 2011, although the coverage is limited before the beginning of 2013.

The CSDB provides comprehensive information about each security. Such information includes, although it is not limited to, international security identification numbers, outstanding amount, currency denomination, security type (e.g., zero

---

7. Source: ECB, https://www.ecb.europa.eu/mopo/implement/omt/html/index.en.html.

coupon bond, perpetuity, etc.), maturity date, issue date, yield to maturity, a set of issuer identifiers (name, legal entity identifier code, bank identifier code, monetary financial institution ID at the ECB), issuers' sector (government, NFC, monetary financial institution, insurance company, pension fund, asset management firm, or other financial institution), and issuer's industry.

We perform an extensive series of cleaning and filtering exercises on the initial CSDB data to improve its quality and make them suitable for our econometric analysis. The detailed procedure is available upon request.

We are interested in euro-denominated bonds issued by NFCs. The reason to exclude foreign-currency-denominated bonds is that their issuance may be related to foreign-exchange expectations that reasonably change at the time of major monetary-policy announcements. Including foreign-currency-denominated securities would therefore make our results difficult to interpret. As for the choice to focus on NFCs, although under the CSPP, bonds issued by non-bank financial institutions may be eligible, we choose to eliminate all financial corporations from the set of securities that we consider. The reason is that QE directly affects both the investment opportunities and the financing opportunities of these institutions, thus preventing us from disentangling whether their issuance has changed in order to accommodate different investment strategies or in order to take advantage of new financing opportunities.

We group euro-denominated debt securities issued by NFCs into two main categories: eligible bond and non-eligible bonds. We define a security as eligible at the end of a given month the ECB accepts it as collateral at that date. The CSDB pro-

vides information about whether each instrument, in any given month, is accepted by the ECB as collateral for its refinancing operations. However, the list of eligible securities is published daily on the ECB's website.

We then use information about the credit rating of the issuers. We include ratings from the four ECB-recognized rating agencies (S&P, Fitch, Moody's and DBRS). We then assign a numerical value to each rating (1 for AAA, 2 for AA+, etc.) and compute the median rating for each security, as a summary statistic of its credit quality.

We then download daily bond prices from Datastream. Prices refer to the end of the day. Although price data are not available for all bonds, we obtain price data for 1,541 bonds, covering 72% of the outstanding amount of bonds issued by NFCs in the period between November 2015 and June 2016. For these bonds, we also collected the bid-ask spreads from Bloomberg.

2.6 Details of the Analysis of Bond Prices and Issuance

In this section, we provide the details of the econometric analysis of the effect of the CSPP announcement on bond prices and issuance that we use to identify the channels of transmission of QE, which we discuss in section 2.7.

We employ a difference-in-differences approach, augmenting our regressions with a series of fixed effects that take into account a series of determinants of the heterogeneous impact of QE on bond prices and issuance. We disentangle the heterogeneity that comes from across-firm variations as well as from within-firm variations, thus controlling for the particular valuation and financing needs of individual firms at any

given time.

We first discuss the details of the price analysis and then the details of the issuance analysis. All results are discussed and compared in section 2.7.

2.6.1   Price Analysis

Here, we describe the variables of interest and the regression specification for our cross-sectional difference-in-differences identification of the transmission channels of monetary policy. Both are chosen in a way that allows us to compare securities differing on a relevant characteristic of interest, while controlling for a series of other factors.

Variables of Interest

The dependent variables of interest are the log prices and the log returns of bonds issued by NFCs in the euro area. We indicate with P_{it} the price of bond i on day t, and its day t log return is given by $\log P_{it} - \log P_{it-1}$.

The explanatory variables of interest are the interaction between security-level (or firm-level) variables and time dummies for the CSPP announcement. As we discuss in section 2.6.1, we choose the explanatory variables and their interaction because a statistical association between the dependent variables of interest and such interactions would be evidence of a particular transmission channel and its importance relative to others.

The bond-level explanatory variables of interest are related to the eligibility of the bond as a collateral, its rating, and its past performance. In particular, we define

the following variables:

- Elig_i, an indicator for whether security i is accepted as collateral by the ECB at the beginning of the sample period;

- Distr_i, an indicator for whether security i was distressed before the announcement of the CSPP;

- $\mathrm{MildDistr}_i$, an indicator for whether security i was experiencing a form of mild distress before the announcement of the CSPP;

- Illiq_i, an indicator for whether the security has a high bid-ask spread.

Securities are defined as distressed or mildly distressed on the basis of their performance in the months before the announcement. For each security, we compute the log-price change from December 10, 2015 (three months before the announcement) to February 11, 2016 (four weeks before the announcement). We then sort securities into deciles according to their price change. We classify the securities in the first decile (the worst performing ones) as distressed, and classify the securities in the second decile (the second worst performing ones) as mildly distressed.

Securities are instead defined as illiquid ($\mathrm{Illiq} = 1$) on the basis of their bid-ask spread, relative to their price. For each security, we compute the average spread-to-price ratio in the period between December 10, 2015, and February 11, 2016. We then sort bonds into quintiles and define as illiquid those securities in the highest quintile.

In terms of transmission channels, Elig distinguishes securities on the basis of their exposure to the scarcity channel, and Distr and MildDistr distinguish securities

on the basis of their exposure to a risk channel. We introduce Illiq as a control.

Table 2.1 in Appendix 2.C reports the correlation between the eligibility indicator of the bond, whether it is investment grade (BBB- or better), indicators for the distress state of the bond, and its illiquidity. The table uses the sample of bonds that we use in our empirical analysis. Although the eligibility is correlated with the investment-grade status and the distress indicators in a intuitive way, the overlap is far from perfect. In a regression setting, this imperfect correlation allows to identify which force is dominant in explaining the cross-sectional response of bond prices to the CSPP announcement.

We also group securities into eight maturity bins and use them to construct maturity-day fixed effects, which partial out movements in the term structure. The maturity bins are the following: (i) up to 6 months, (ii) 6 months to 1 year, (iii) 1 to 2 years, (iv) 2 to 5 years, (v) 5 to 10 years, (vi) 10 to 20 years, (vii) 20 to 30 years, and (viii) longer than 30 years.

When looking at returns, we are interested in knowing whether returns, on the day of the announcement, are larger for a set of securities than for others. However, liquidity, for some bonds, may be an issue, and we therefore consider the possibility that part of the effect may be observed also on the day following the announcement. When looking at prices, we ask whether prices of some set of bonds experience a persistent increase relative to some other set of bonds, following the announcement. Therefore, the time dummies to be used in our analysis are:

- EventDay, an indicator that takes the value of 1 on the day of the CSPP announcement;

161

- DayAfter, an indicator that takes the value of 1 on the day following the CSPP announcement day;

- Post, an indicator that takes the value of 1 from the day of the CSPP announcement onward.

We use the first two time dummies for the analysis of returns, and the third one for the analysis of prices.

We also construct firm-level explanatory variables that are analogous to the bond-level ones:

- $EligF_i$, an indicator for whether security i is issued by a firm that issues bonds that are accepted as collateral by the ECB at the beginning of the sample period;

- $DistrF_i$, an indicator for whether security i is issued by a firm that was distressed before the announcement of the CSPP;

- $MildDistrF_i$, an indicator for whether security i is issued by a firm that was experiencing a form of mild distress before the announcement of the CSPP.

Similarly to the Distr and MildDistr dummies, DistrF and MildDistrF are constructed based on the past price performance. However, this time, we aggregate the total value of the bonds issued by each firm and group firms into deciles according to the log change of their total bond value. Firms in the first decile are classified as distressed, and firms in the second decile are classified as mildly distressed.

Similarly to the bond-level variable, EligF captures the exposure of firms to the scarcity channel, DistrF and MildDistrF capture the exposure of firms to a risk channel, and we introduce InvGrF as a control in some regression specifications.

Empirical Strategy

We use the variables introduced in section 2.6.1 in a linear regression setting. Coefficients on the interactions of bond-level (or firm-level) variables with time dummies are associated with the transmission channels of monetary policy. For example, when we analyze prices, a positive coefficient on the Elig*Post interaction indicates a scarcity channel is present and it more than compensates for a liquidity channel. The coefficients on Distr*Post and MildDistr*Post capture the extent to which market impairment, changes in risk perception, or in expectation of future (risk-adjusted) profitability determines the transmission on unconventional monetary policy.

We begin with the simple possible model to detect the presence of a scarcity channel. For returns, we estimate

$$\log P_{it} - \log P_{it-1} = \beta_1 \text{EventDay}_t \cdot \text{Elig}_i + \beta_2 \text{DayAfter}_t \cdot \text{Elig}_i + \text{Fixed Effects} + u_{it},$$

(2.2)

whereas for log prices, we use

$$\log P_{it} = \beta \text{Post}_t \cdot \text{Elig}_i + \text{Fixed Effects} + u_{it}.$$

(2.3)

To disentangle the possible transmission channels in the cross-section, we ex-

plore augmented models that include interactions of the distress dummies with time dummies.

In particular, for the daily returns, we consider the following regression:

$$\log P_{it} - \log P_{it-1} =$$

$$\beta_1 \text{EventDay}_t \cdot \text{Distr}_i + \gamma_1 \text{EventDay}_t \cdot \text{Distr}_i + \delta_1 \text{EventDay}_t \cdot \text{MildDistr}_i$$

$$+\beta_2 \text{DayAfter}_t \cdot \text{Distr}_i + \gamma_2 \text{DayAfter}_t \cdot \text{Distr}_i + \delta_2 \text{DayAfter}_t \cdot \text{MildDistr}_i \quad (2.4)$$

$$+\text{Fixed Effects} + u_{it},$$

whereas for log prices, we estimate

$$\log P_{it} = \beta \text{Post}_t \cdot \text{Elig}_i + \gamma \text{Post}_t \cdot \text{Distr}_i + \delta \text{Post}_t \cdot \text{MildDistr}_i + \text{Fixed Effects} + u_{it}.$$
$$(2.5)$$

All specifications include a security fixed effect that takes into account the unobserved heterogeneity across securities (and hence issuers) and of the initial price level, a fixed effect for the coupon-type-day combination that takes into account that different coupon types (fixed, floating, zero, etc.) may be related to heterogeneous responses to monetary-policy announcements, and a maturity-day fixed effect to partial out effects coming from the movement in the term structure of the interest rates. We also show results when rating-day fixed effects are excluded or included. These should partially absorb some of the effects that are associated with expectations of higher chances of rating upgrade and with changing risk premia. However, note that the current credit rating of a bond may not properly capture some risk components.

In the baseline specification, we use country-industry-day fixed effects to control for the heterogeneous exposure of different countries and industries to unconventional monetary policy and to control for demand conditions. However, this specification does not allow us to disentangle whether the effects that we observe are due to heterogeneous price impact across bonds within a given firm or across firms themselves.

To test whether the effect derives from heterogeneity within firms, we also estimate a model using firm-day fixed effects. We then compare these results to the estimates of a model that uses country-industry-day fixed effects, but uses firm-level regressors EligF, DistrF, and MildDistrF instead of their bond-level counterparts.

We use a three-month window around the announcement date and consider only securities whose price is available throughout the entire period, thus avoiding concerns related to new issuance and expirations. We also consider only securities with ratings between BBB+ and BB, to ensure the sample of securities has comparable credit quality, while leaving a degree of heterogeneity that is rich enough to identify the desired effects. We then obtain a sample of 391 bonds observed over a period of 131 days. Of these bonds, 70.3% are eligible, 87.0% have investment grade, 8.1% of them experienced financial distress before the CSPP announcement, and 7.6% were in a state of mild distress.

We present results for both non-weighted and weighted regressions, with weights given by the initial outstanding amount of each bond. Standard errors are always clustered in order to allow for unrestricted correlation at the security and country-industry-day level.

2.6.2   Issuance Analysis

We now describe the variables of interest and the regression specification for the difference-in-differences analysis of bond issuance. This analysis mirrors the analysis of bond prices, with some differences due to the less granular and lower-frequency nature of the bond issuance data. However, our strategy remains that of comparing firms and types of issuance that are differently exposed to a channel of transmission of interest, while controlling for a set of other factors.

As before, we study whether the effects that we observe are an across-firm or a within-firm phenomenon, by choosing either country-industry-month or firm-month fixed effects.

Variables of Interest

We are interested in the issuance of bonds by individual NFCs in the euro area. For each firm, we distinguish between the issuance of securities that are eligible and securities that are not eligible. Therefore, each month t, all issuance observations are indexed by the pair (i, Elig), where i indexes the firm and $\text{Elig} \in \{0, 1\}$ is an indicator for whether the issuance is eligible or not. We denote with I_{it}^{Elig} the net issuance of bonds by firm i in month t, with Elig indicating the eligibility class of such issue. Net issuance is expressed in nominal (i.e., face value) terms and is computed as issuance of new securities plus changes in the outstanding amount of existing securities, including expirations.

To take into account the different size of various firms, we scale issuance by the total outstanding amount of bonds of firm i at the beginning of the sample period,

B_{i0}. Therefore, the LHS variable of interest is

$$\frac{I_{it}^{\text{Elig}}}{B_{i0}},$$

which we call scaled net issuance. Tables 2.14 and 2.15 in Appendix 2.D report summary statistics for the net scaled issuance at the firm-eligibility level in the 3 months and 10 months around the announcement, respectively.

We also conduct firm-level analysis in which the variable of interest is the total net issuance of a firm, which we indicate as I_{it} and which is equal to $I_{it}^0 + I_{it}^1$. This variable is also scaled by B_{i0} in our empirical analysis.

As before, we exploit the interaction between issue-level (or firm-level) variables and time dummies for the CSPP announcement in order to detect the presence of transmission channels of monetary policy. We focus in this section on the presence of a scarcity channel and the incentives of firms to favor eligible bonds versus non-eligible bonds. Therefore, the only issue-level explanatory variable of interest is Elig, which, as mentioned before, indexes issues that are eligible for CSPP.

The relevant time dummies are instead

- EventMonth, an indicator that takes the value of 1 in the month of the CSPP announcement;

- MonthAfter, an indicator that takes the value of 1 in the month following the CSPP announcement month;

- NextMonths, an indicator that takes the value of 1 in all the subsequent months included in the sample.

167

We also conduct analyses with firm-level variables. In particular, we use

- EligF$_i$, which tells whether firm i had eligible bonds outstanding at the beginning of the sample period;

- DistrF$_i$, an indicator for whether security i is issued by a firm that was distressed before the announcement of the CSPP;

- MildDistrF$_i$, an indicator for whether security i is issued by a firm that was experiencing a form of mild distress before the announcement of the CSPP.

The variables DistrF and MildDistrF are constructed as in section 2.6.1.

Empirical Strategy

We test for the presence of a scarcity-channel-induced change in the issuance policy by regressing the net scaled issuance on the eligibility-time-dummy interactions and fixed effects. A positive coefficient on, for example, Elig*NextMonths indicate that, following the CSPP announcement, firms have increased the issuance of eligible bonds more than the issuance of non-eligible bonds, even after a two-month lag. If such an effect is observed, we are led to conclude, on the basis of the previous results, that a scarcity channel is present, but its relative price effects are neutralized by the action of firms that adapt their issuance policy to take advantage of the investors' higher marginal appetite for eligible securities.

We therefore estimate a linear model in the form

$$
\frac{I_{it}^{\text{Elig}}}{B_{i0}} = \alpha \text{Elig}
$$

$$
+ \beta_1 \text{EventMonth}_t \cdot \text{Elig} + \beta_2 \text{MonthAfter}_t \cdot \text{Elig} + \beta_3 \text{NextMonths}_t \cdot \text{Elig}
$$

$$
+ \text{Fixed Effects} + u_{it}^{\text{Elig}}.
$$

(2.6)

We specify the model with firm fixed effects or with firm-eligibility fixed effects (in the latter case, the coefficient α is not identified). In the baseline specification, we include country-industry-month fixed effects.

To assess whether the results that we obtain from the baseline model are the results of within- or across-firm heterogeneity, we estimate a model that includes firm-month fixed effects, thus looking only at firms that issue both eligible and non-eligible bonds.

We compare the results from the model with firm-month fixed effects with the results of a model that uses country-industry-month fixed effects but employs the firm-level eligibility dummy EligF, instead of the issuance-level dummy Elig. In this latter model, we aggregate the total issuance at the firm-month level, so that, letting $I_{it} = I_{it}^0 + I_{it}^1$, the model for across-firm heterogeneity is in the form

$$
\frac{I_{it}}{B_{i0}} = \alpha \text{EligF}_i
$$

$$
+ \beta_1 \text{EventMonth}_t \cdot \text{EligF}_i + \beta_2 \text{MonthAfter}_t \cdot \text{EligF}_i + \beta_3 \text{NextMonths}_t \cdot \text{EligF}_i
$$

$$
+ \text{Fixed Effects} + u_{it},
$$

and we use firm fixed effects and country-industry-month fixed effects.

169

We also estimate a version of model (2.6) that includes interactions between the distress dummies DistrF and MildDistrF and the time dummies. We also include firm fixed effects or firm-eligibility fixed effects and country-industry-month fixed effects.

We perform these regression analyses employing a three- and a 10-month window around the announcement. In both cases, we consider only the firm-eligibility pairs (i, Elig) that are observed at the beginning of the sample period. In other words, we consider the issuance of eligible (ineligible) bonds by firm i only if firm i has a positive amount of eligible (ineligible) bonds outstanding at the beginning of the sample period. We do so to avoid dealing with issues related to the endogeneity of the eligibility of the issuer. For example, we would like to avoid concerns about issuers receiving a rating upgrade because of a more accomodative monetary policy, thus allowing their issuance to be eligible only after (and because of) the announcement.

When a three-month window is considered, the dataset contains 2,476 issuers, of which 182 are eligible (i.e., had eligible securities at the beginning of the sample period) and 2,383 issue non-eligible securities, meaning 89 of them issued both eligible and non-eligible securities. When a 10-month windows is used, 2,396 issuers are present in the dataset, of which 177 are eligible, 2,301 issue non-eligible securities, and 82 issue both eligible and non-eligible bonds.

All regressions are weighted by the firms' total outstanding amount of bonds at the beginning of the sample period, B_{i0}. We show results for three different levels of clustering of the standard errors: one that allows only for heteroschedasticity, one that allows also for correlation within firm-eligibility pairs, and one that allows for

170

correlations among all observation belonging to the same country-industry-eligibility group.

2.7 Results on the Transmission Channels of QE

We now discuss the evidence that our cross-sectional analysis provides on the transmission channels of QE. In particular, we separately discuss the available evidence, or lack thereof, of a scarcity, liquidity, and risk channel.

All tables are in the Appendix, with Appendix 2.C containing all regression tables for the analysis of the price effect of the CSPP announcement, and Appendix 2.D containing all regression tables for the issuance.

2.7.1 Scarcity Channel

We find evidence of a scarcity channel because firms increased the issuance of eligible securities relative to non-eligible securities. We find no evidence of an increase in the price of eligible securities relative to non-eligible securities. Moreover, the relative cost of borrowing for eligible and non-eligible firms seems to be unaffected by this channel, because eligible firms did not issue more bonds than non-eligible firms in response to the announcement. Our results indicate, therefore, that although investors may have a preferred habitat, firms, being natural short sellers of their own debt, adjust the relative issuance of eligible and non-eligible debt, although total issuance remains unaffected by a scarcity channel.

On the day of the announcement, eligible assets experienced lower returns, which resulted in a persistently lower price change relative to non-eligible assets in the three

171

months following the announcement. Focusing on column (4) of Table 2.2, we observe that, in the two days following the announcement, eligible bonds experienced returns that, on average, were 70 basis points lower than non-eligible bonds, even after controlling for any effect that is coming from the country and industry of the issuer, shifts in the term structure, and changes in the market valuation of credit ratings. In Table 2.3, we observe that the differential price change tended to be persistent, even if its statistical significance depends on the weighting and the inclusion of rating-day fixed effects.

Most of the heterogeneity in the price change happens across firm variation, with no heterogeneity in the price change across eligible and non-eligible bonds within a given firm. Moreover, the across-firm effect tends to be more strongly persistent. Indeed, when we control for firm-day fixed effects in Table 2.4, we observe essentially no heterogeneity in the price change of eligible and non-eligible securities. In the three months following the announcement, eligible securities appreciate relative to non-eligible securities of the same issuers, as shown in Table 2.5, although the difference is not statistically significant. However, as shown in Tables 2.6 and 2.7, eligible firms strongly under-performed non-eligible firms both on the announcement day and in the subsequent three months in terms of their overall debt valuation. Focusing on columns (4) in both tables, we observe eligible firms experiencing returns of 87 basis points below non-eligible firms in the two days after the announcement, with the return gap widening in the subsequent three months to 182 basis points.

As for the issuance, after the announcement, we observe an increase in eligible issuance compared to non-eligible issuance. This increase can be observed in Table

172

2.16 where, if we focus on column (6), we estimate that the eligible monthly issuance increased by 1.87% relative to non-eligible monthly issuance in the three months after the announcement. Given that the total amount of bonds outstanding for the firms in the sample was €888.4bn at the end of November 2015, the monthly net issuance of eligible securities increased by €16.6bn more than the monthly net issuance of non-eligible securities.

In contrast to what we find for prices, the heterogeneity in the issuance can be primarily observed within firms, because we find no change in the total issuance of eligible firms relative to non-eligible firms. This finding indicates that a scarcity channel is unlikely to lower the cost of capital of firms. As we see in column (4) of Table 2.18, firms that issued both eligible and non-eligible bonds increased the eligible issuance by 2.53% in the three months after the announcement, compared to the previous three. Given that their total outstanding amount of bonds was €451bn in November 2015, these firms increased the issuance of eligible bonds relative to non-eligible ones by €11.4bn each month. The fact that estimates for the coefficient of interest are larger when we control for firm-month fixed-effects demand for financing, suggests the presence of a negative correlation between the firm-level monetary policy shock and firm-level shocks to demand for financing. Therefore, following Khwaja and Mian (2008) and Jiménez et al. (2014), we can be reasonably confident tha, if we conduct firm-level analysis, we would not be overestimating the impact of the CSPP announcement. Backed by these consideration, in Table 2.20, we compare the total issuance of eligible firms with the total issuance of non-eligible firms. We find a small negative impact of firm eligibility on total issuance, although not statistically

significant. This result indicates that eligible firms did not borrow more than non-eligible firms, thus suggesting that the scarcity channel was not effective at lowering the relative cost of borrowing.

The results of the within-firm analysis provides a clarification regarding the mechanism behind our findings. Firms issuing both eligible and non-eligible bonds are likely unconstrained, and therefore they can freely adjust their liability structure in response to changes in the market condition. If they act as neoclassical investors, even if habitat investors are present, the only possible equilibrium after QE is the one where firms adjust the relative quantity of eligible and non-eligible issuance up to the point where the habitat investors' valuation of the different bond categories is aligned with the firms', which coincides with the fundamental valuation. Therefore, we should observe no heterogeneity in the price change of bonds issued by these firms. However, quantities should adjust so that issuance of eligible bonds increases, which is exactly what we observe in the data for this set of firms.

2.7.2   Liquidity Channel

We find no evidence of a liquidity channel. Indeed, if a liquidity channel was a driver of the lower returns of eligible securities on the day of the announcement, we should observe an analogous fact when the PSPP was announced, because it involved monthly purchases for €60bn each month.

We therefore repeat the same analysis we conducted for the CSPP also for the PSPP and find no heterogeneity across eligible and non-eligible securities. Regression results are available in the online Appendix. In Appendix 2.E, we simply report time-

174

series plots of corporate bond prices around the PSPP announcement as a simple illustration of the lack of any heterogeneous effect of unconventional monetary policy.

We can therefore comfortably reject the hypothesis that a liquidity channel is responsible for the transmission of QE to the euro area bond prices and issuance.

Moreover, this result allows us to rule out the explanation that eligible and non-eligible bonds simply load differently on monetary-policy factors. Indeed, if the heterogeneity we observe were simply due to the fact that non-eligible bonds load more on monetary policy factors, we would observe them increase in price more than eligible ones also after the PSPP announcement.

2.7.3 Risk Channel

We find that the risk channel is the main driver of the transmission of QE to bond prices. Specifically, the price of distressed bonds increased more on the day of the announcement than the price of other bonds. However, distressed firms did not increase issuance, relative to non-distressed firms. Because of this latter result, we hesitate to claim that a risk channel operates through risk prices. If QE decreased risk prices, then riskier firms would face lower costs of borrowing and less binding financing constraints, and they would therefore issue more bonds. Because we fail to find evidence of this outcome, we are inclined to instead support the idea that the riskiness of the underlying asset decreased. Other explanations for the lack of increased issuance for these firms may include the fact that distressed firms are financially constrained or that they may be already operating below capacity.

On the announcement day, distressed bonds experienced consistently higher re-

turns than other firms, which also resulted in persistently different price changes. If we look at Table 2.8, we can observe that distressed and mildly distressed bonds experienced significantly higher returns than other bonds. This differential effect can be observed also within eligible and non-eligible firms, with magnitudes that, for distressed bonds, range between 1.5% and 2% over a two-day period. The difference persists in the following months, as we see in Table 2.9. However, in this longer period, distressed bonds issued by non-eligible firms seem to be the ones experiencing the largest increase in valuation, with a price increase of 4.3% when we control for illiquidity. By contrast, the valuation of distressed bonds issued by eligible firms increase by 2.6%.

The role of distress in the transmission of QE to prices can also be observed across bonds issued by the same firm. In Tables 2.10 and 2.11, we control for firm-day fixed effects and still find a large and statistically significant effect of past performance on the price increase of bonds following the announcement. The results also hold when we use firm-level distress indicators, as in Tables 2.12 and 2.13. However, we cannot establish a comparison between eligible and non-eligible firms due to the strong collinearity of the eligibility and distress dummy variables at firm level.

If we look at the issuance of distressed firms, we find no evidence that it has increased. In Tables 2.22 and 2.24, we include dummies for the distress status of the issuer, thus augmenting the regression in Tables 2.16 and 2.20. Instead of observing distressed firms issuing more than non-distressed ones, the coefficients on the Distr and MildDistr interactions indicate that distressed firms did not issue more than non-distressed ones.

176

Overall, our results suggest that a risk channel is a key component of the transmission of QE to bond prices and a crucial determinant in how different firms benefit from unconventional monetary policy.

2.8 Conclusions

We provided a cross-sectional analysis of the effects of quantitative easing on corporate bond prices and issuance. We exploit the fact that non-eligible bonds experienced a period of distress in the months leading to the CSPP announcement in order to clearly disentangle a scarcity channel and a risk channel.

We found that although a scarcity channel seems to be present, its effect can be mainly observed within firms and its price effects are neutralized by firms changing their mix of issued bonds in order to satisfy investors' demand.

As far as prices are concerned, the most powerful channels seem to be related to changes in the quantity and/or price of risk. This channel mainly manifests itself in heterogeneous responses of bond prices across firms. Firms whose bonds experienced the largest price decline in the months preceding the QE announcement were the ones that experienced the largest increase in their debt valuation after the announcement of the program. However, no heterogeneous effect on the issuance can be detected in the cross section of firms on the basis of past performance.

Although the latter two facts seem to be mutually inconsistent, they can be reconciled in a model where the private sector tend to take excessive risk compared to a socially optimal allocation, as in Stein (2012), Greenwood et al. (2016), and Woodford (2016). This behavior may appear if, for the example, the private sector

177

has incentives to issue money-like securities that are subject to runs but does not internalize the social cost of early liquidation. If QE corrects these incentives by satiating the economy's demand for money, we may at the same time observe a reduction in the economy-wide risk and a little, or even negative, change in total bond issuance. The reason is that absent QE, the privately chosen level of issuance would be characterized by an inefficient mix of liabilities and/or excessive issuance.

Although developing a full theoretical model was beyond the scope of this paper, our results should constitute the empirical background for any model trying to explain the channels through which QE affects market prices, financing decisions, and, ultimately, real economic activity.

Appendix 2.A   Further Details of the ECB's Asset-Purchase Programs

In this appendix, we provide a brief overview of the ECB's asset-purchase programs. Because we already discussed the CSPP in detail in section 2.4.2, here we mainly focus on the remaining programs: the third Covered Bond Purchase Program (CBPP3), the Asset-Back Securities Purchase Program (ABSPP), and the Public Sector Purchase Program (PSPP)

Details on the exact size of each individual program were not usually disclosed ex ante. However, when the PSPP was announced on January 22, 2015, the ECB disclosed the intention to buy a total amount of €60bn each month of securities covered by the ABSPP, CBPP3, and PSPP programs. This total amount was increased to €80bn when the CSPP was announced, on March 10, 2016. However, the ECB publishes its security holdings and monthly purchases at the end of every month, aggregated by program.

The duration of the APPs was first announced on January 22, 2015, when the ECB stated its intention to carry on asset purchases until September 2016. The duration was then prolonged to March 2017, following a press conference on December 3, 2015, when the ECB also announced it would keep reinvesting the principal payments on the securities purchased under the APPs for as long as necessary. On January 19, 2017, the ECB further prolonged the extension of the APPs to December 2017, deciding, however, to reduce the monthly purchases to €60bn from April 2017.

Finally, the Eurosystem makes available for security lending all the assets pur-

chased under the PSPP and CSPP and publishes a list of the available ISINs, because, given the large scale of the APPs, the ECB wanted to set up adequate facilities to prevent episodes of collateral scarcity.

Below, we describe the program-specific details, including the eligibility criteria for securities to be covered by each program and the rationale that the ECB put forward for these programs.

The Asset-Back Securities Purchase Program (ABSPP)

Under the ABSPP, the ECB carried on purchase of asset-backed securities with the stated purpose of reducing banks' funding risk and stimulating lending by facilitating the sales of pools of loans in the secondary market.

The ABSPP was announced on September 4, 2014, and the detailed technical modalities were release on October 2, 2014. Purchases began on November 21, 2014, and as of May 2017, the ECB was holding €23,653mn in securities purchased under the ABSPP program.

The Third Covered Bond Purchase Program (CBPP3)

The CBPP3 is a credit-easing measure aimed at facilitating the extension of credit to the real economy through the purchase of covered bonds issued by euro-area monetary financial institutions (MFIs). Indeed, a covered bond is a debt instruments, usually issued by a credit institution, that is backed by a pool of assets capable of covering the claims arising from the bond until its maturity. Moreover, the assets

backing the covered bond are used as collateral by the covered bondholders in case of default of the issuer.

The CBPP3 is the third in a series of CBPPs that was initiated in July 2009. The CBPP3 was announced on September 4, 2014, and the detailed technical modalities were released on October 2, 2014. Purchases began on October 20, 2014, and as of May 2017, the ECB was holding €219,927mn in securities purchased under the CBPP3 program.

The Public Sector Purchase Program (PSPP)

The PSPP is the largest of APPs and encompasses securities issued by the euro area's public sector including central governments, recognized agencies, local governments, and international and multilateral organizations[8].

This program was undertaken with the purpose of providing further monetary accommodation in a low rate environment, anchor medium- and long-term inflation expectations, and provide support to the forward guidance of the ECB.

Furthermore, to be eligible for any APP, a security must have a yield to maturity that exceeds the deposit facility rate, in order to avoid arbitrage.

The PSPP was announced on January 22, 2015, and the detailed technical modalities were released on the same day. Purchases began on March 9, 2015, and as of May 2017, the ECB was holding €1,563bn in securities purchased under the PSPP program.

The PSPP imposes restrictions on the eligibility requirements of securities. In

---

8. The latter two account for about 10% of the PSPP purchases.

particular, public sector securities must be accepted as collateral at the ECB (see section 2.4.2 for details) and, initially, they must have had a residual maturity at the time of purchase between 2 and 30 years. Moreover, the yield to maturity of securities must be above the deposit facility rate at the time of the purchase, although some exceptions have been allowed starting from January 2, 2017. From the same date, the lower limit on the residual maturity has been decreased to one year.

Appendix 2.B Plots

2.B.1   NFC Bond Market
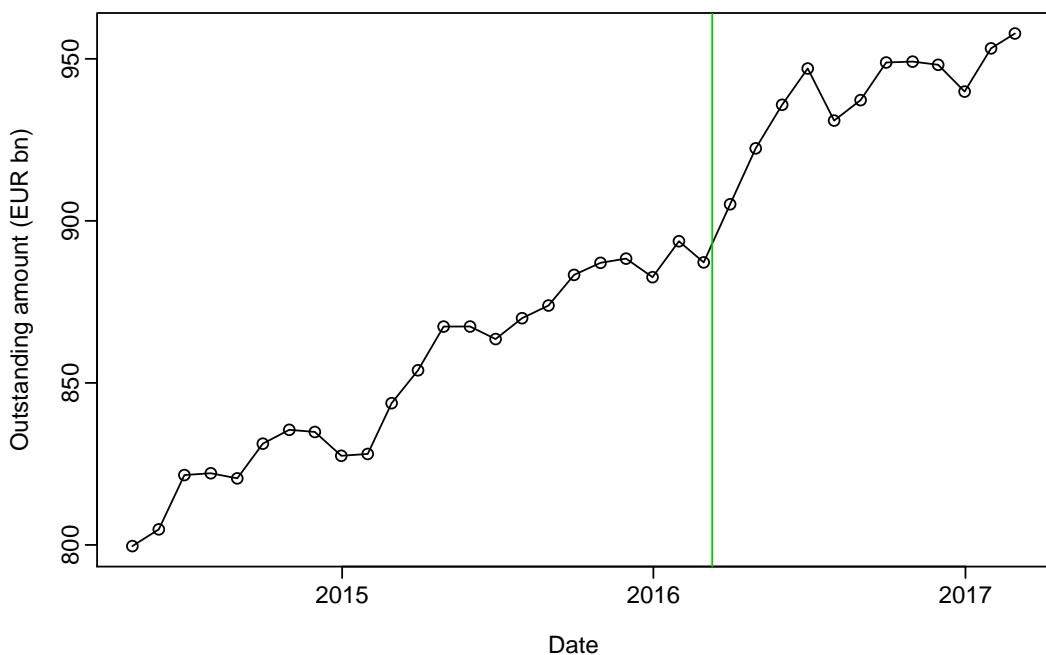
Figure 2.1: Outstanding amount of euro-denominated bonds issued by non-financial corporations in the Euro-area. The vertical line marks the announcement of the CSPP (March 10, 2016).
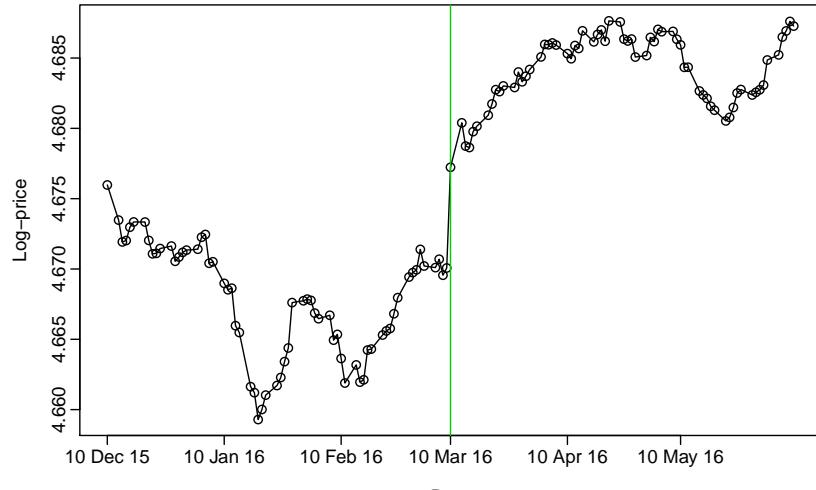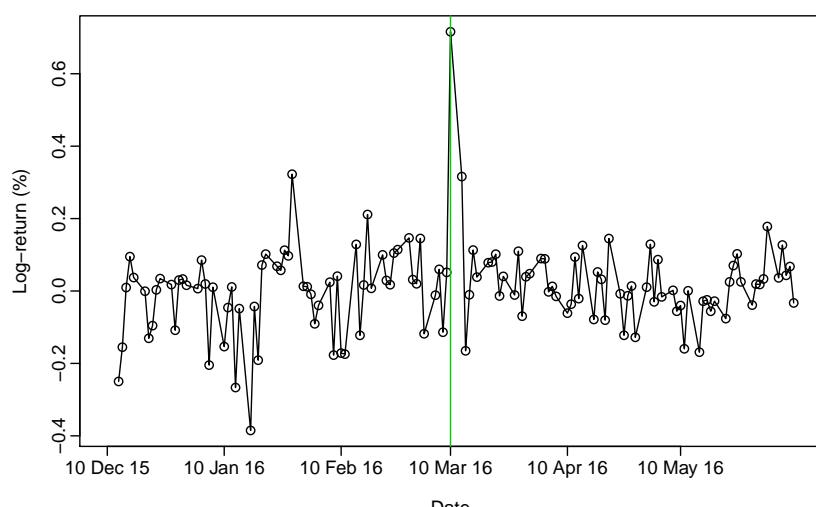
(a) Log-prices

(b) Log-returns

Figure 2.2: Weighted average of log-prices and log-returns of bonds issued by Euro-area NFC with bond rating between BBB+ and BB. The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the CSPP (March 10, 2016).
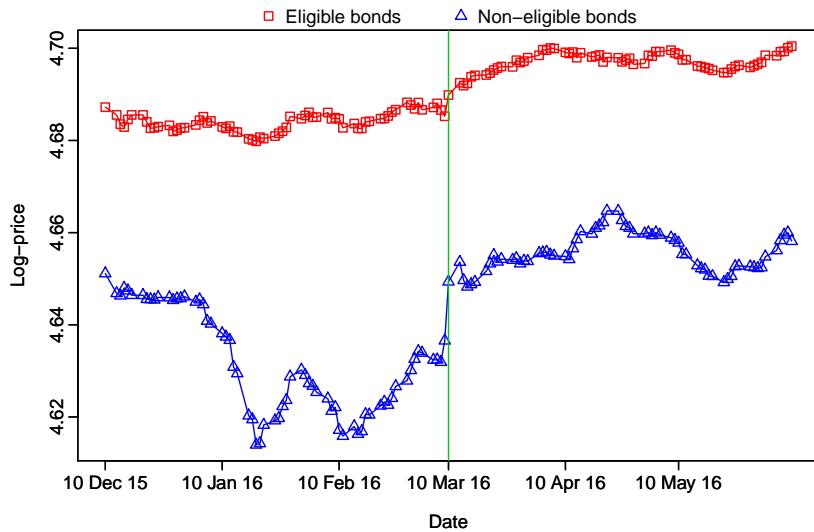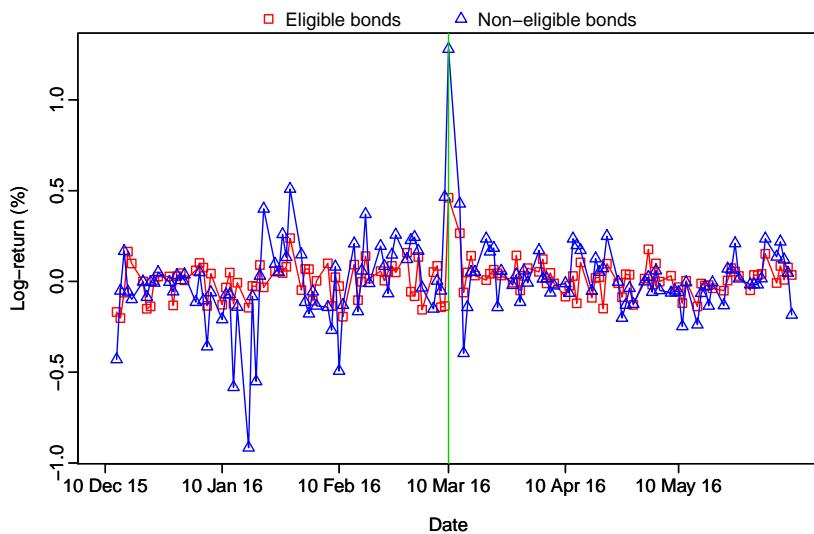
(a) Log-prices

(b) Log-returns

Figure 2.3: Weighted average of log-prices and log-returns of eligible and non-eligible bonds issued by Euro-area NFC with bond rating between BBB+ and BB. The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the CSPP (March 10, 2016).
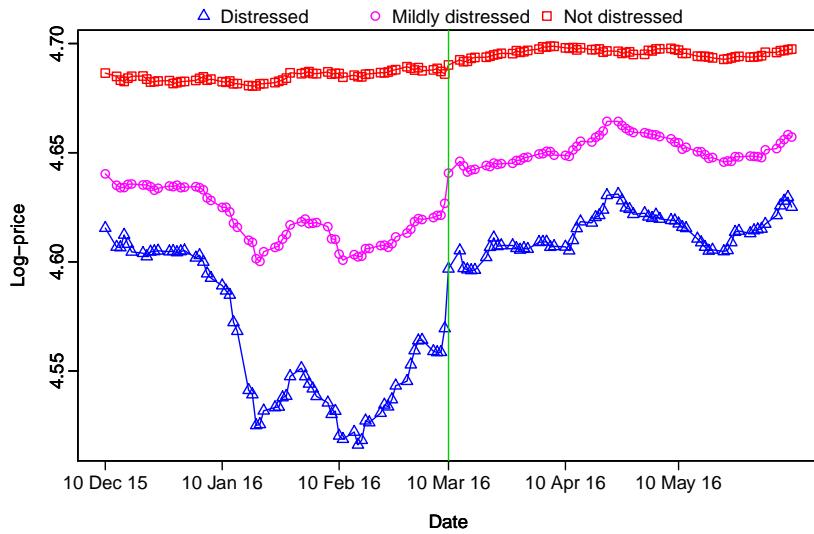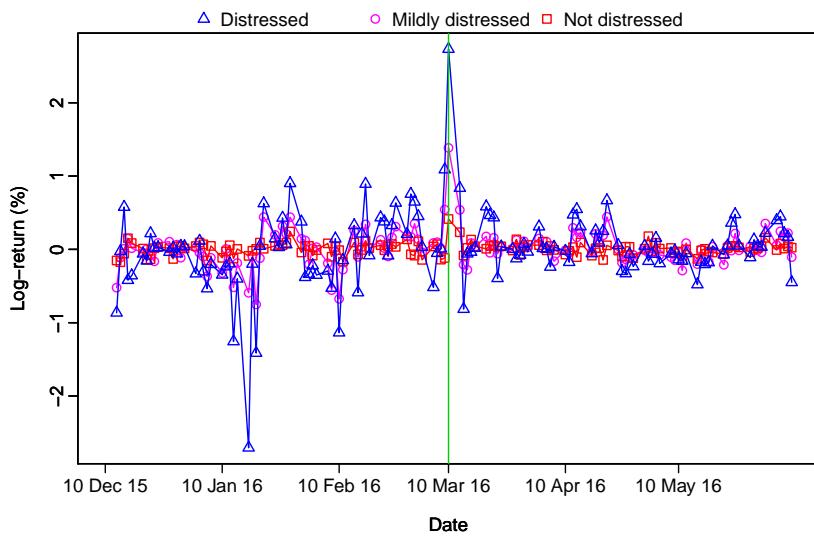
(a) Log-prices

(b) Log-returns

Figure 2.4: Weighted average of log-prices and log-returns of bonds issued by Euro-area NFC with bond rating between BBB+ and BB. Bonds are classified as distressed, mildly distressed and not distressed. The distress state is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the CSPP (March 10, 2016).
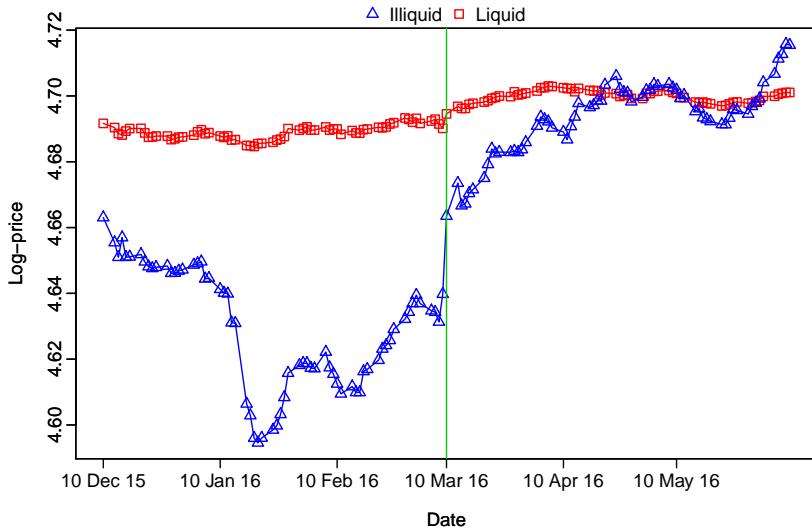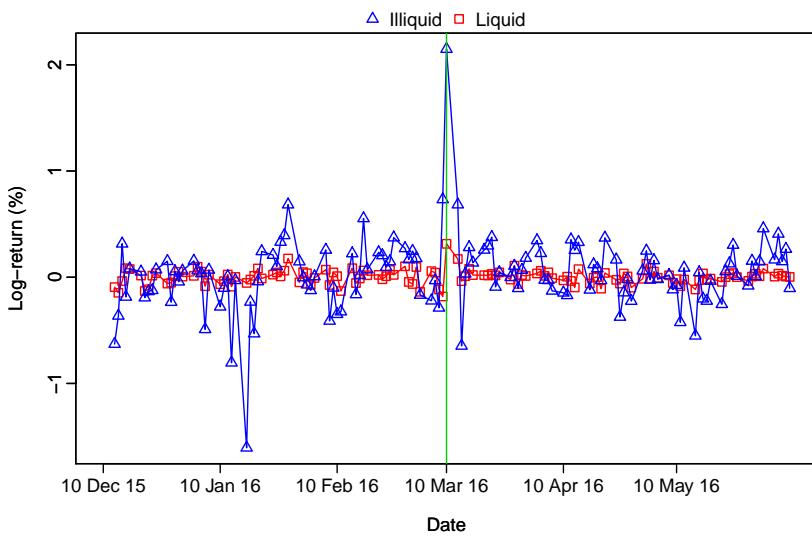
(a) Log-prices

(b) Log-returns

Figure 2.5: Weighted average of log-prices and log-returns of bonds issued by Euro-area NFC with bond rating between BBB+ and BB. Bonds are classified as liquid or illiquid. Illiquidity is defined on the basis of the average bid-ask spread in the months before the announcement (see Section 2.6.1 for details). The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the CSPP (March 10, 2016).

2.B.3   Issuance around the CSPP Announcement
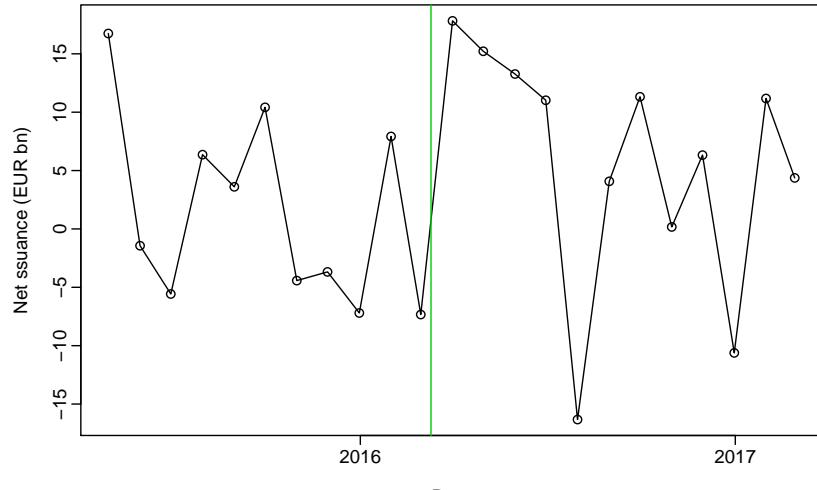
(a) All Bonds
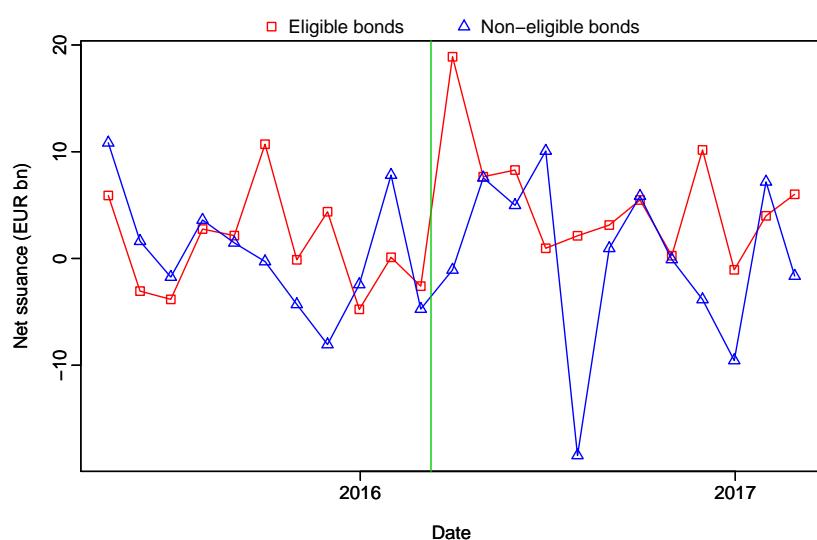
(b) Bonds by Eligibility

Figure 2.6: Net issuance of euro-denominated bonds issued by non-financial corporations in the Euro-area. Here, bonds are defined as eligible if they can be used as collateral for the ECB's refinancing operations. The vertical line marks the announcement of the CSPP (March 10, 2016).

187
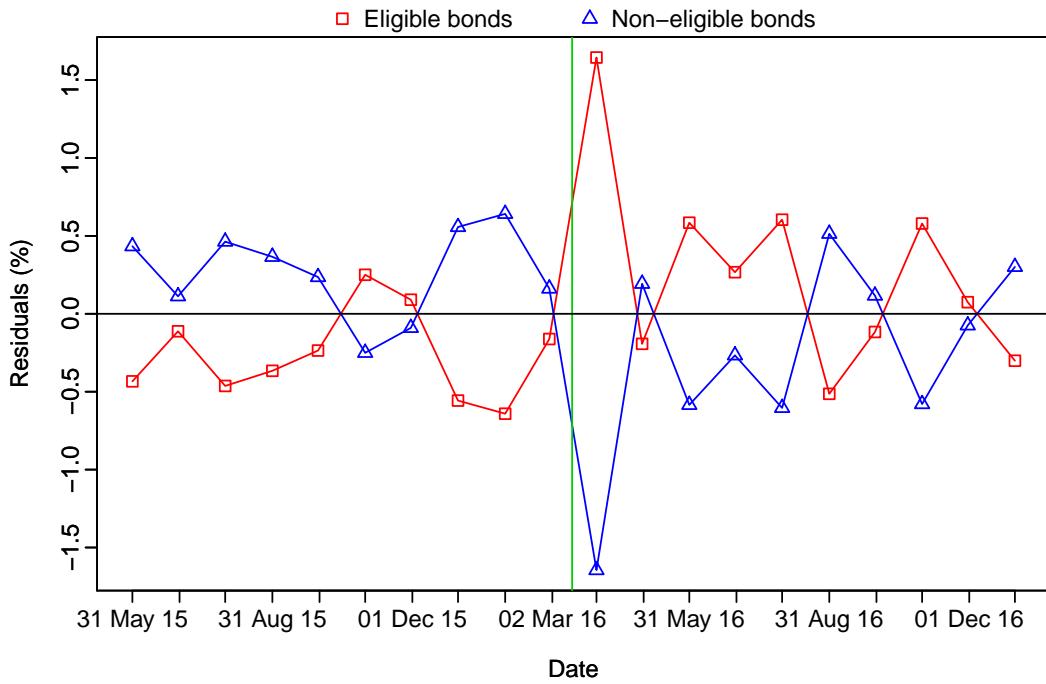
Figure 2.7: Residuals of a project of firm's scaled net issuance by eligibility ($I_{it}^{\text{Elig}}/B_{i0}$) on firm-month and firm-eligibility fixed effects. Net issuance I_{it}^{Elig} represents the net issuance of eligible or non eligible bonds (Elig = 1 or Elig = 0, respectively) by firm i in month t. The vertical line marks the announcement of the CSPP (March 10, 2016).

188

Appendix 2.C   Effect of CSPP on Bond Prices – Tables

Table 2.1: Correlation between bond-level indicator variables. Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. InvGr = 1 if the bond has median rating above BBB- at the beginning of the sample period. Distr and MildDistr take value 1 if the bond is, respectively, in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). Finally, Illiq = 1 if the bond has a high bid-ask spread in the period before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB.

	Elig	InvGr	Distr	MildDistr	Illiq
Elig	1	0.347	-0.296	-0.170	-0.081
InvGr	0.347	1	-0.161	-0.288	0.145
Distr	-0.296	-0.161	1	-0.086	0.315
MildDistr	-0.170	-0.288	-0.086	1	0.097
Illiq	-0.081	0.145	0.315	0.097	1

Table 2.2: Regression for the announcement effect of CSPP on daily bond returns. Event-Day = 1 on March 10, 2016 (announcement day), while DayAfter = 1 on the following day. Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. The coefficient in the row "Compound effect" is the sum of the coefficients in the previous two rows. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | |
	(1)	(2)	(3)	(4)
EventDay*Elig	−0.329	−0.491**	−0.347**	−0.505***
	(0.205)	(0.238)	(0.175)	(0.190)
DayAfter*Elig	−0.238**	−0.271**	−0.198***	−0.195**
	(0.093)	(0.123)	(0.073)	(0.083)
Compound effect	-0.567**	-0.762***	-0.545***	-0.700***
	(0.227)	(0.262)	(0.181)	(0.194)
Country-industry-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	49,010	49,010	49,010	49,010
R^2	0.375	0.434	0.420	0.489

| *Notes:* | $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$ |

Table 2.3: Regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-price (%) | | | |
	(1)	(2)	(3)	(4)
Post*Elig	-1.180***	-1.071**	-0.785***	-0.618
	(0.390)	(0.538)	(0.283)	(0.398)
Country-industry-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	49,387	49,387	49,387	49,387
R^2	0.983	0.984	0.986	0.987

Notes:	$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.4: Regression for the announcement effect of CSPP on daily bond returns. Event-Day = 1 on March 10, 2016 (announcement day), while DayAfter = 1 on the following day. Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. The coefficient in the row "Compound effect" is the sum of the coefficients in the previous two rows. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses firm-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | |
	(1)	(2)	(3)	(4)
EventDay*Elig	−0.152	−0.318	−0.053	−0.105
	(0.191)	(0.300)	(0.148)	(0.175)
DayAfter*Elig	0.059	0.117*	0.031	0.065
	(0.059)	(0.065)	(0.051)	(0.040)
Compound effect	-0.093	-0.201	-0.022	-0.040
	(0.206)	(0.288)	(0.160)	(0.174)
Firm-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	45,370	45,370	45,370	45,370
R^2	0.559	0.601	0.576	0.627
Notes:	$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$			

Table 2.5: Regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses firm-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

	Log-price (%)			
	(1)	(2)	(3)	(4)
Post*Elig	0.558	1.336	0.555	1.168*
	(0.484)	(0.888)	(0.396)	(0.627)
Firm-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	45,719	45,719	45,719	45,719
R^2	0.992	0.993	0.993	0.994
Notes:	$^*p \le .10$; $^{**}p \le .05$; $^{***}p \le .01$			

193

Table 2.6: Regression for the announcement effect of CSPP on daily bond returns. Event-Day = 1 on March 10, 2016 (announcement day), while DayAfter = 1 on the following day. EligF = 1 if the bond is issued by a firm having bonds that are eligible as collateral at the beginning of the sample period. The coefficient in the row "Compound effect" is the sum of the coefficients in the previous two rows. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | |
	(1)	(2)	(3)	(4)
EventDay*EligF	−0.510	−1.244***	−0.126	−0.694***
	(0.320)	(0.433)	(0.208)	(0.246)
DayAfter*EligF	−0.265	−0.557***	−0.186	−0.180*
	(0.177)	(0.212)	(0.173)	(0.094)
Compound effect	-0.775**	-1.801***	-0.312	-0.874***
	(0.382)	(0.500)	(0.242)	(0.180)
Country-industry-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	49,010	49,010	49,010	49,010
R^2	0.538	0.692	0.776	0.947
Notes:	$^*p \le .10;\ ^{**}p \le .05;\ ^{***}p \le .01$			

Table 2.7: Regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). EligF = 1 if the bond is issued by a firm having bonds that are eligible as collateral at the beginning of the sample period. The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

	Log-price(%)			
	(1)	(2)	(3)	(4)
Post*EligF	-2.266***	-4.538***	-1.344***	-1.816***
	(0.468)	(0.562)	(0.307)	(0.314)
Country-industry-day FE	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes
Rating-day FE	No	No	Yes	Yes
Weighted	No	Yes	No	Yes
Observations	49,387	49,387	49,387	49,387
R^2	1.000	1.000	1.000	1.000

Notes:	$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.8: Unweighted regression for the announcement effect of CSPP on daily bond returns. EventDay = 1 on March 10, 2016 (announcement day), while DayAfter = 1 on the following day. Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. Distr and MildDistr take value 1 if the bond is, respectively, in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | | | |
| | All firms | | Eligible firms | | Ineligible firms | |
	(1)	(2)	(3)	(4)	(5)	(6)
EventDay*Elig	−0.146	−0.139	−0.185	−0.190		
	(0.101)	(0.099)	(0.122)	(0.122)		
EventDay*Distr	1.481***	1.382***	1.285***	1.224***	1.694***	1.368***
	(0.325)	(0.308)	(0.434)	(0.418)	(0.556)	(0.496)
EventDay*MildDistr	0.498***	0.462**	0.267	0.245	0.534	1.122
	(0.188)	(0.182)	(0.178)	(0.171)	(0.968)	(1.218)
EventDay*Illiq		0.291*		0.177		1.285
		(0.170)		(0.151)		(0.789)
DayAfter*Elig	−0.129**	−0.127**	−0.072	−0.074		
	(0.062)	(0.064)	(0.064)	(0.065)		
DayAfter*Distr	0.551***	0.523***	0.626***	0.604***	0.271	0.149
	(0.151)	(0.152)	(0.214)	(0.202)	(0.472)	(0.314)
DayAfter*MildDistr	0.406**	0.396**	0.505**	0.497**	0.011	0.232
	(0.175)	(0.177)	(0.223)	(0.228)	(0.651)	(0.784)
DayAfter*Illiq		0.083		0.066		0.482
		(0.146)		(0.115)		(0.689)
Elig two-day effect	-0.276***	-0.266***	-0.257*	-0.264**		
	(0.099)	(0.098)	(0.132)	(0.131)		
Distr two-day effect	2.032***	1.905***	1.911***	1.828***	1.965***	1.517**
	(0.343)	(0.332)	(0.448)	(0.431)	(0.763)	(0.601)
MildDistr two-day effect	0.904***	0.858***	0.772***	0.741***	0.546	1.354
	(0.228)	(0.226)	(0.259)	(0.258)	(0.986)	(1.066)
Illiq two-day effect		0.374*		0.243		1.767**
		(0.218)		(0.188)		(0.734)
Country-industry-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,010	49,010	42,900	42,900	6,110	6,110
R^2	0.409	0.409	0.481	0.481	0.751	0.753

Notes: $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.9: Unweighted regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. Distr and MildDistr take value 1 if the bond is, respectively, in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-price (%) | | | | | |
| | All firms | | Eligible firms | | Ineligible firms | |
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Elig	−0.425*	−0.395*	0.060	0.027		
	(0.239)	(0.233)	(0.349)	(0.344)		
Post*Distr	2.953***	2.449***	3.080***	2.679**	5.224***	4.287***
	(0.938)	(0.941)	(1.107)	(1.105)	(1.190)	(1.023)
Post*MildDistr	1.710***	1.531***	1.718***	1.572***	0.790	2.832***
	(0.450)	(0.439)	(0.524)	(0.514)	(1.129)	(0.937)
Post*Illiq		1.496***		1.180***		3.786***
		(0.305)		(0.292)		(1.273)
Country-industry-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,387	49,387	43,230	43,230	6,157	6,157
R^2	0.986	0.986	0.990	0.990	0.993	0.994

Notes: $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.10: Unweighted regression for the announcement effect of CSPP on daily bond returns. EventDay $= 1$ on March 10, 2016 (announcement day), while DayAfter $= 1$ on the following day. Elig $= 1$ if the bond is eligible as collateral at the beginning of the sample period. Distr and MildDistr take value 1 if the bond is, respectively, in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses firm-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | | | |
| | All firms | | Eligible firms | | Ineligible firms | |
	(1)	(2)	(3)	(4)	(5)	(6)
EventDay*Elig	−0.061	−0.071	−0.055	−0.062		
	(0.118)	(0.116)	(0.120)	(0.118)		
EventDay*Distr	1.444***	1.391***	1.353***	1.321***	1.377***	0.949**
	(0.421)	(0.430)	(0.466)	(0.473)	(0.490)	(0.466)
EventDay*MildDistr	0.669***	0.678***	0.571*	0.570**	0.661	1.273
	(0.251)	(0.249)	(0.295)	(0.290)	(0.926)	(1.297)
EventDay*Illiq		0.293*		0.184		1.398
		(0.172)		(0.153)		(1.146)
DayAfter*Elig	0.025	0.019	0.029	0.025		
	(0.049)	(0.045)	(0.047)	(0.045)		
DayAfter*Distr	0.396**	0.361**	0.426*	0.409*	0.196	−0.043
	(0.197)	(0.184)	(0.235)	(0.228)	(0.658)	(0.467)
DayAfter*MildDistr	0.435*	0.441*	0.477	0.477	0.115	0.457
	(0.238)	(0.241)	(0.324)	(0.327)	(0.906)	(0.964)
DayAfter*Illiq		0.192*		0.101		0.780
		(0.105)		(0.089)		(0.780)
Elig two-day effect	-0.036	-0.052	-0.026	-0.036		
	(0.134)	(0.129)	(0.134)	(0.130)		
Distr two-day effect	1.840***	1.753***	1.779***	1.730***	1.573*	0.907
	(0.425)	(0.431)	(0.477)	(0.483)	(0.915)	(0.700)
MildDistr two-day effect	1.104***	1.120***	1.048***	1.047***	0.776	1.730
	(0.273)	(0.276)	(0.361)	(0.361)	(1.089)	(1.209)
Illiq two-day effect		0.485**		0.286*		2.178**
		(0.198)		(0.169)		(1.023)
Firm-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45,370	45,370	40,820	40,820	4,550	4,550
R^2	0.571	0.571	0.547	0.547	0.731	0.733

Notes: $^{*}p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

198

Table 2.11: Regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). Elig = 1 if the bond is eligible as collateral at the beginning of the sample period. Distr and MildDistr take value 1 if the bond is, respectively, in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bond in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses firm-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-price (%) | | | | | |
| | All firms | | Eligible firms | | Ineligible firms | |
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Elig	0.622	0.568	0.632	0.587		
	(0.443)	(0.427)	(0.425)	(0.413)		
Post*Distr	3.228***	2.972***	2.758**	2.568**	5.198***	4.038***
	(1.003)	(0.980)	(1.101)	(1.103)	(1.354)	(1.236)
Post*MildDistr	1.329***	1.381***	0.998**	0.995**	0.798	2.911***
	(0.352)	(0.357)	(0.470)	(0.471)	(1.265)	(1.004)
Post*Illiq		1.474***		1.142***		4.084***
		(0.286)		(0.243)		(1.514)
Firm-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45,719	45,719	41,134	41,134	4,585	4,585
R^2	0.994	0.994	0.994	0.994	0.989	0.992

| Notes: | $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$ |

Table 2.12: Unweighted regression for the announcement effect of CSPP on daily bond returns. EventDay = 1 on March 10, 2016 (announcement day), while DayAfter = 1 on the following day. EligF = 1 if the bond is issued by a Firm having bonds that are eligible as collateral at the beginning of the sample period. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

| | Log-return (%) | | | | | |
| | All firms | | Eligible firms | | Ineligible firms | |
	(1)	(2)	(3)	(4)	(5)	(6)
EventDay*EligF	−0.069	−0.068				
	(0.195)	(0.194)				
EventDay*DistrF	1.788***	1.787***				
	(0.195)	(0.198)				
EventDay*MildDistrF	0.496*	0.495*	0.221	0.226		
	(0.258)	(0.260)	(0.152)	(0.150)		
EventDay*Illiq		0.008		−0.028		0.057
		(0.076)		(0.074)		(0.044)
DayAfter*EligF	−0.101	−0.108				
	(0.127)	(0.124)				
DayAfter*DistrF	0.516**	0.531**				
	(0.209)	(0.217)				
DayAfter*MildDistrF	0.479**	0.489**	0.544***	0.550***		
	(0.195)	(0.190)	(0.135)	(0.136)		
DayAfter*Illiq		−0.081		−0.033		−0.248*
		(0.068)		(0.031)		(0.130)
EligF two-day effect	-0.170	-0.176				
	(0.224)	(0.220)				
DistrF two-day effect	2.304***	2.318***				
	(0.306)	(0.317)				
MildDistrF two-day effect	0.975***	0.985***	0.764***	0.776***		
	(0.357)	(0.356)	(0.226)	(0.226)		
Illiq two-day effect		-0.073		-0.061		-0.191
		(0.104)		(0.089)		(0.139)
Country-industry-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,010	49,010	42,900	42,900	6,110	6,110
R^2	0.583	0.583	0.776	0.776	0.962	0.962

Notes: $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.13: Unweighted regression for the announcement effect of CSPP on daily bond prices. Post = 1 on and after March 10, 2016 (announcement day). EligF = 1 if the bond is issued by a Firm having bonds that are eligible as collateral at the beginning of the sample period. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). The sample period covers the three months prior and post the announcement. The sample contains only securities that are continuously traded throughout the sample period and that have median rating between BBB+ and BB. This table uses country-industry-day fixed effects. Standard errors are clustered at country-industry-day and security level. Standard errors are in parentheses.

	Log-price (%)					
	All firms		Eligible firms		Ineligible firms	
	(1)	(2)	(3)	(4)	(5)	(6)
Post*EligF	−0.996	−0.985				
	(0.796)	(0.801)				
Post*DistrF	3.320*	3.295*				
	(1.910)	(1.903)				
Post*MildDistrF	1.399***	1.381***	1.255***	1.202***		
	(0.377)	(0.375)	(0.370)	(0.381)		
Post*Illiq		0.137		0.275		−0.346
		(0.190)		(0.217)		(0.466)
Country-Industry-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Security FE	Yes	Yes	Yes	Yes	Yes	Yes
Coupon.Type-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Rating-day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,387	49,387	43,230	43,230	6,157	6,157
R^2	1.000	1.000	1.000	1.000	1.000	1.000

Notes: $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Appendix 2.D Effect of CSPP on Bond Issuance – Tables

Table 2.14: Summary statistics for scaled net issuance by firm-eligibility pair in a 3 month window around the CSPP announcement.

Net Issuance (%)		Summary Statistics							
Eligibility	Months	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
All	All	15,390	8.617	37.183	−362.371	0.000	0.000	0.000	1,365.654
Non-eligible	All	14,298	8.930	38.325	−362.371	0.000	0.000	0.000	1,365.654
Eligible	All	1,092	4.514	15.330	−32.645	0.000	0.000	0.971	210.526
All	Before CSPP	7,695	8.286	34.067	−362.371	0.000	0.000	0.000	950.000
All	Post CSPP	7,695	8.948	40.057	−168.421	0.000	0.000	0.000	1,365.654
Non-eligible	Before CSPP	7,149	8.645	35.153	−362.371	0.000	0.000	0.000	950.000
Non-eligible	Post CSPP	7,149	9.215	41.255	−168.421	0.000	0.000	0.000	1,365.654
Eligible	Before CSPP	546	3.582	12.356	−14.324	0.000	0.000	0.000	84.211
Eligible	Post CSPP	546	5.445	17.778	−32.645	0.000	0.000	2.127	210.526

Table 2.15: Summary statistics for scaled net issuance by firm-eligibility pair in a 10 month window around the CSPP announcement.

Net Issuance by Eligibility (%)		Summary Statistics							
Eligibility	Months	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
All	All	49,560	8.914	160.007	−615.385	0.000	0.000	0.000	33,839.500
Non-eligible	All	46,020	9.306	165.984	−615.385	0.000	0.000	0.000	33,839.500
Eligible	All	3,540	3.818	15.710	−62.888	0.000	0.000	0.871	513.636
All	Before CSPP	24,780	8.906	55.958	−300.000	0.000	0.000	0.000	5,555.556
All	Post CSPP	24,780	8.923	219.258	−615.385	0.000	0.000	0.000	33,839.500
Non-eligible	Before CSPP	23,010	9.281	57.833	−300.000	0.000	0.000	0.000	5,555.556
Non-eligible	Post CSPP	23,010	9.331	227.503	−615.385	0.000	0.000	0.000	33,839.500
Eligible	Before CSPP	1,770	4.021	18.253	−35.714	0.000	0.000	0.809	513.636
Eligible	Post CSPP	1,770	3.615	12.670	−62.888	0.000	0.000	0.953	150.000

Table 2.16: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 3 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Net Issuance by Eligibility (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	0.937**		0.937**		0.937**	
	(0.392)		(0.427)		(0.392)	
EventMonth*Elig	3.632*	3.632**	3.632*	3.632*	3.632*	3.632*
	(2.028)	(1.729)	(2.006)	(2.014)	(2.108)	(2.117)
MonthAfter*Elig	0.225	0.225	0.225	0.225	0.225	0.225
	(0.801)	(0.825)	(0.759)	(0.762)	(0.493)	(0.495)
NextMonths*Elig	1.865**	1.865**	1.865**	1.865**	1.865***	1.865***
	(0.840)	(0.809)	(0.774)	(0.777)	(0.612)	(0.615)
Country-industry-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	No	Yes	No	Yes	No
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	13,919	13,919	13,919	13,919	13,919	13,919
R^2	0.484	0.542	0.484	0.542	0.484	0.542

Notes:

$^{*}p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.17: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 10 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Net Issuance by Eligibility (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	1.001***		1.001***		1.001**	
	(0.213)		(0.389)		(0.416)	
EventMonth*Elig	3.016	3.016	3.016	3.016	3.016	3.016
	(1.973)	(1.876)	(1.963)	(1.965)	(2.034)	(2.036)
MonthAfter*Elig	−0.285	−0.285	−0.285	−0.285	−0.285	−0.285
	(0.707)	(0.648)	(0.652)	(0.653)	(0.400)	(0.401)
NextMonths*Elig	0.753**	0.754**	0.753**	0.754**	0.753***	0.754***
	(0.317)	(0.299)	(0.304)	(0.304)	(0.234)	(0.234)
Country-industry-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	No	Yes	No	Yes	No
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	42,038	42,038	42,038	42,038	42,038	42,038
R^2	0.374	0.395	0.374	0.395	0.374	0.395

Notes:

$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.18: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 3 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. This table uses firm-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Net Issuance by Eligibility (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	0.701		0.701		0.701	
	(0.469)		(0.519)		(0.538)	
EventMonth*Elig	4.032	4.032	4.032	4.032	4.032	4.032
	(3.059)	(2.780)	(2.993)	(3.277)	(3.039)	(3.327)
MonthAfter*Elig	0.571	0.571	0.571	0.571	0.571	0.571
	(0.871)	(1.017)	(0.700)	(0.767)	(0.720)	(0.788)
NextMonths*Elig	2.532**	2.532**	2.532***	2.532**	2.532***	2.532***
	(1.076)	(1.091)	(0.931)	(1.019)	(0.886)	(0.971)
Firm-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	1,068	1,068	1,068	1,068	1,068	1,068
R^2	0.502	0.645	0.502	0.645	0.502	0.645

Notes:

$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.19: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 10 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. This table uses firm-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Net Issuance by Eligibility (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	0.913***		0.913*		0.913	
	(0.250)		(0.529)		(0.591)	
EventMonth*Elig	3.816	3.816	3.816	3.816	3.816	3.816
	(3.158)	(3.054)	(3.141)	(3.221)	(3.195)	(3.278)
MonthAfter*Elig	0.141	0.141	0.141	0.141	0.141	0.141
	(0.830)	(0.699)	(0.693)	(0.711)	(0.642)	(0.659)
NextMonths*Elig	0.820**	0.820**	0.820**	0.820**	0.820**	0.820**
	(0.376)	(0.335)	(0.358)	(0.367)	(0.336)	(0.344)
Firm-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	3,280	3,280	3,280	3,280	3,280	3,280
R^2	0.496	0.644	0.496	0.644	0.496	0.644

Notes:

$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.20: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on total bond issuance by firms using a 3 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. EligF = 1 if the firm has bonds that are eligible as collateral at the beginning of the sample period. This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Total Net Issuance (%)		
	(1)	(2)	(3)
EventMonth*EligF	3.978**	3.978*	3.978
	(2.021)	(2.273)	(3.267)
MonthAfter*EligF	−2.293	−2.293	−2.293**
	(1.735)	(1.974)	(1.088)
NextMonths*EligF	−0.666	−0.666	−0.666
	(1.448)	(1.523)	(1.433)
Country-industry-month FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Firm clustered SE	No	Yes	No
Country-industry-eligibility clustered SE	No	No	Yes
Observations	13,373	13,373	13,373
R^2	0.587	0.587	0.587
Notes:	$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$		

Table 2.21: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on total bond issuance by firms using a 10 month window around the announcement. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. EligF = 1 if the firm has bonds that are eligible as collateral at the beginning of the sample period. This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Total Net Issuance (%)		
	(1)	(2)	(3)
EventMonth*EligF	3.655*	3.655*	3.655
	(1.983)	(2.046)	(2.864)
MonthAfter*EligF	−1.310	−1.310	−1.310
	(1.348)	(1.405)	(0.831)
NextMonths*EligF	−0.137	−0.137	−0.137
	(0.539)	(0.567)	(0.548)
Country-industry-month FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Firm clustered SE	No	Yes	No
Country-industry-eligibility clustered SE	No	No	Yes
Observations	40,360	40,360	40,360
R^2	0.408	0.408	0.408
Notes:	$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$		

Table 2.22: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 3 month window around the announcement with firm distress indicators. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

| | Net Issuance by Eligibility (%) | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	0.943**		0.943**		0.943**	
	(0.402)		(0.433)		(0.399)	
EventMonth*Elig	3.593	3.593*	3.593*	3.593	3.593	3.593
	(2.198)	(1.900)	(2.173)	(2.216)	(2.294)	(2.339)
EventMonth*DistrF	0.839	0.839	0.839	0.839	0.839	0.839
	(1.697)	(1.652)	(1.682)	(1.715)	(1.760)	(1.795)
EventMonth*MildDistrF	−0.288	−0.288	−0.288	−0.288	−0.288	−0.288
	(2.897)	(2.491)	(2.621)	(2.673)	(3.060)	(3.121)
MonthAfter*Elig	0.089	0.089	0.089	0.089	0.089	0.089
	(0.844)	(0.877)	(0.786)	(0.801)	(0.510)	(0.520)
MonthAfter*DistrF	−1.676	−1.676	−1.676	−1.676	−1.676	−1.676
	(1.464)	(1.437)	(1.362)	(1.389)	(1.319)	(1.345)
MonthAfter*MildDistrF	−4.717	−4.717	−4.717	−4.717	−4.717	−4.717
	(3.690)	(3.080)	(3.559)	(3.630)	(3.732)	(3.805)
NextMonths*Elig	2.016**	2.016**	2.016**	2.016**	2.016***	2.016***
	(0.873)	(0.855)	(0.790)	(0.806)	(0.618)	(0.630)
NextMonths*DistrF	−1.059	−1.059	−1.059	−1.059	−1.059	−1.059
	(1.815)	(1.804)	(1.894)	(1.931)	(1.700)	(1.733)
NextMonths*MildDistrF	3.261	3.261	3.261	3.261	3.261	3.261
	(3.891)	(3.828)	(4.602)	(4.693)	(4.495)	(4.584)
Country-industry-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	No	Yes	No	Yes	No
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	2,877	2,877	2,877	2,877	2,877	2,877
R^2	0.392	0.489	0.392	0.489	0.392	0.489

Notes: $^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.23: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 10 month window around the announcement with firm distress indicators. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. Elig = 1 if the issuance is eligible as collateral. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

	Net Issuance by Eligibility (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elig	1.000***		1.000**		1.000**	
	(0.223)		(0.408)		(0.435)	
EventMonth*Elig	3.039	3.039	3.039	3.039	3.039	3.039
	(2.163)	(2.063)	(2.155)	(2.165)	(2.248)	(2.259)
EventMonth*DistrF	0.627	0.627	0.627	0.627	0.627	0.627
	(1.367)	(1.309)	(1.330)	(1.337)	(1.340)	(1.347)
EventMonth*MildDistrF	1.575	1.575	1.575	1.575	1.575	1.575
	(2.028)	(1.779)	(1.662)	(1.670)	(1.644)	(1.652)
MonthAfter*Elig	−0.454	−0.454	−0.454	−0.454	−0.454	−0.454
	(0.769)	(0.704)	(0.708)	(0.711)	(0.436)	(0.438)
MonthAfter*DistrF	−1.903*	−1.903*	−1.903*	−1.903*	−1.903**	−1.903**
	(1.096)	(1.049)	(1.037)	(1.042)	(0.948)	(0.953)
MonthAfter*MildDistrF	−2.900	−2.900	−2.900	−2.900	−2.900	−2.900
	(2.366)	(2.093)	(2.062)	(2.073)	(2.698)	(2.712)
NextMonths*Elig	0.750**	0.750**	0.750**	0.750**	0.750***	0.750***
	(0.338)	(0.318)	(0.322)	(0.323)	(0.253)	(0.254)
NextMonths*DistrF	−0.886	−0.886	−0.886	−0.886	−0.886	−0.886
	(0.909)	(0.892)	(1.004)	(1.009)	(0.999)	(1.004)
NextMonths*MildDistrF	0.759	0.759	0.759	0.759	0.759	0.759
	(0.988)	(0.961)	(1.228)	(1.235)	(1.129)	(1.135)
Country-industry-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	No	Yes	No	Yes	No
Firm-eligibility FE	No	Yes	No	Yes	No	Yes
Firm-eligibility clustered SE	No	No	Yes	Yes	No	No
Country-industry-eligibility clustered SE	No	No	No	No	Yes	Yes
Observations	9,154	9,154	9,154	9,154	9,154	9,154
R^2	0.478	0.517	0.478	0.517	0.478	0.517

Notes: $^{*}p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$

Table 2.24: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 3 month window around the announcement with firm distress and eligibility indicators. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. EligF = 1 if the firm has bonds that are eligible as collateral at the beginning of the sample period. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

| | Total Net Issuance (%) | | |
	(1)	(2)	(3)
EventMonth*EligF	4.692	4.692	4.692
	(2.983)	(3.391)	(4.924)
EventMonth*DistrF	2.850	2.850	2.850
	(2.736)	(2.993)	(4.165)
EventMonth*MildDistrF	0.026	0.026	0.026
	(3.417)	(3.962)	(4.152)
MonthAfter*EligF	-4.172^*	-4.172	-4.172^{**}
	(2.372)	(2.783)	(1.708)
MonthAfter*DistrF	-5.497^{**}	-5.497^{**}	-5.497^{***}
	(2.305)	(2.498)	(1.942)
MonthAfter*MildDistrF	-7.490	-7.490	-7.490
	(5.067)	(6.515)	(7.386)
NextMonths*EligF	-1.018	-1.018	-1.018
	(1.973)	(2.077)	(1.870)
NextMonths*DistrF	-3.595	-3.595	-3.595^*
	(2.510)	(2.701)	(2.147)
NextMonths*MildDistrF	2.199	2.199	2.199
	(4.594)	(5.763)	(5.844)
Country-industry-month FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Firm clustered SE	No	Yes	No
Country-industry-eligibility clustered SE	No	No	Yes
Observations	2,349	2,349	2,349
R^2	0.573	0.573	0.573
Notes:		$^*p \leq .10$; $^{**}p \leq .05$; $^{***}p \leq .01$	

Table 2.25: Regression for the announcement effect of the CSPP (announced on March 10, 2016) on bond issuance using a 10 month window around the announcement with firm distress and eligibility indicators. EventMonth = 1 on March 2016, MonthAfter = 1 on April 2016, NextMonths = 1 in the subsequent months. EligF = 1 if the firm has bonds that are eligible as collateral at the beginning of the sample period. DistrF and MildDistrF take value 1 if the bond is issued, respectively, by a firm that was in a state of distress or mild distress before the announcement. Distress is defined on the basis of the price performance of the bonds of the firm in the months before the announcement (see Section 2.6.1 for details). This table uses country-industry-month fixed effects. Standard errors are corrected for heteroschedasticity or otherwise clustered at the specified level. Standard errors are in parentheses. All regressions are weighted by the outstanding amount of bonds of the issuer at the beginning of the sample period.

|	Total Net Issuance (%)		
	(1)	(2)	(3)
EventMonth*EligF	4.522	4.522	4.522
	(2.817)	(2.909)	(4.182)
EventMonth*DistrF	2.977	2.977	2.977
	(2.480)	(2.620)	(3.534)
EventMonth*MildDistrF	2.227	2.227	2.227
	(2.240)	(2.157)	(2.423)
MonthAfter*EligF	−2.917	−2.917	−2.917***
	(1.839)	(1.949)	(1.111)
MonthAfter*DistrF	−4.638**	−4.638**	−4.638***
	(1.835)	(1.903)	(1.366)
MonthAfter*MildDistrF	−4.831*	−4.831	−4.831
	(2.852)	(3.096)	(3.969)
NextMonths*EligF	−0.702	−0.702	−0.702
	(0.712)	(0.823)	(0.825)
NextMonths*DistrF	−2.120*	−2.120	−2.120
	(1.160)	(1.387)	(1.361)
NextMonths*MildDistrF	0.431	0.431	0.431
	(1.244)	(1.601)	(1.641)
Country-industry-month FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Firm clustered SE	No	Yes	No
Country-industry-eligibility clustered SE	No	No	Yes
Observations	7,522	7,522	7,522
R^2	0.548	0.548	0.548

Notes: $^*p \leq .10;$ $^{**}p \leq .05;$ $^{***}p \leq .01$

Appendix 2.E Effects of the PSPP Announcement
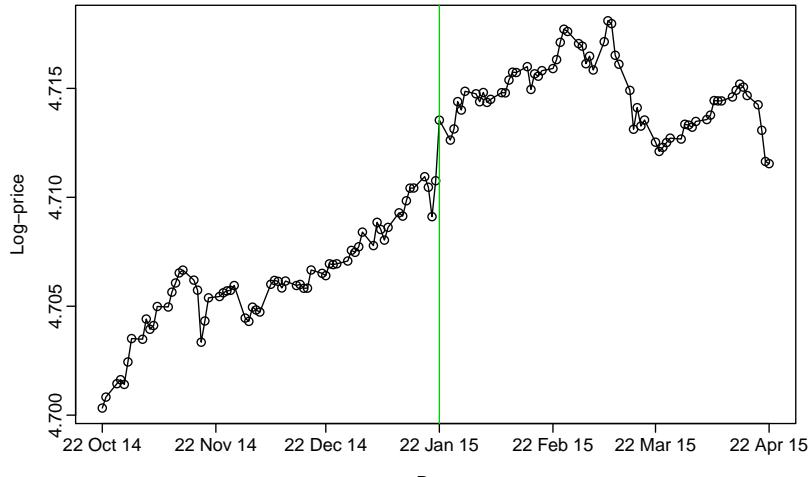

(a) Log-prices
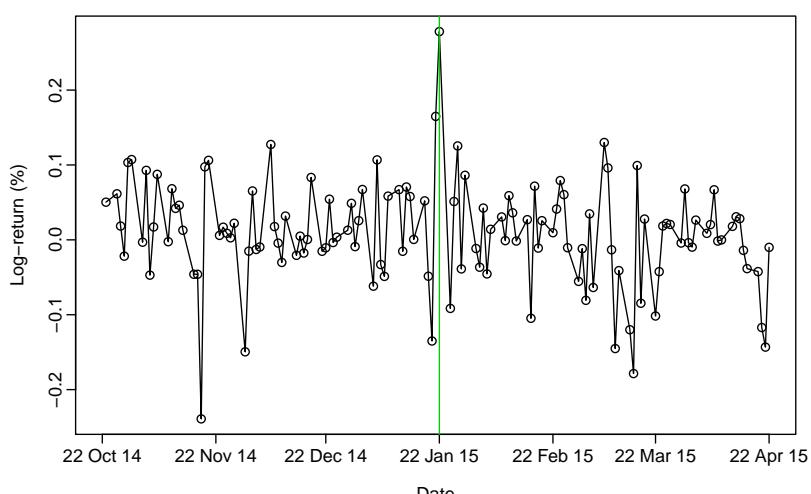
(b) Log-returns

Figure 2.8: Weighted average of log-prices and log-returns of bonds issued by Euro-area NFC with bond rating between BBB+ and BB. The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the PSPP (January 22, 2015).
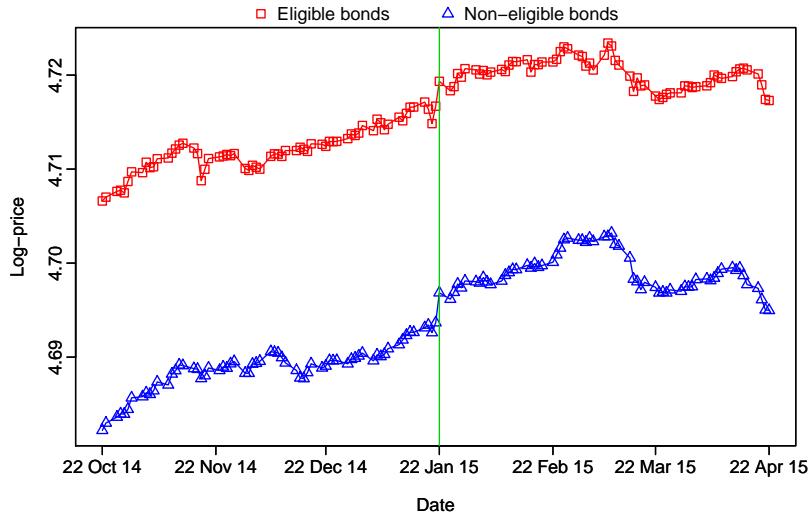
(a) log-prices
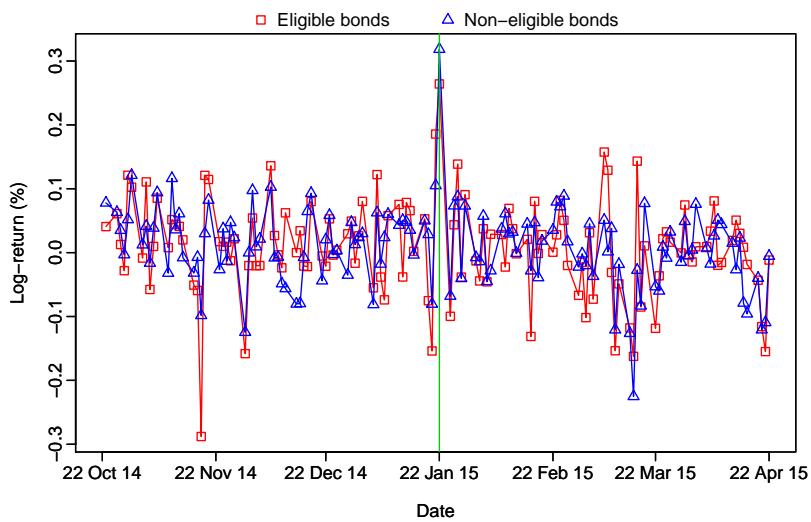
(b) Log-returns

Figure 2.9: Weighted average of log-prices and log-returns of eligible and non-eligible bonds issued by Euro-area NFC with bond rating between BBB+ and BB. The weights are given by the initial outstanding amount of the bond. Bonds in this figure have been traded for the entire sample period. The vertical line marks the announcement of the PSPP (January 22, 2015).

214

References

Abidi, N., Miquel Flores, I., and Eterovic, N. A. (2017). Who benefits from the CSPP? A Difference-in-discontinuities approach.

Altavilla, C., Carboni, G., and Motto, R. (2015). Asset purchase programmes and financial markets: Lessons from the Euro Area. Technical report, European Central bank.

Andrade, P., Breckenfelder, J., De Fiore, F., Karadi, P., and Tristani, O. (2016). The ECB's asset purchase programme: An early assessment. Technical report, European Central Bank.

Baker, M., Greenwood, R., and Wurgler, J. (2003). The maturity of debt issues and predictable variation in bond returns. *Journal of Financial Economics*, 70(2):261–291.

Baker, M. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance*, 55(5):2219–2257.

Barley, R. (2016, February 3). Corporate bonds: Lingering distortion or looming disaster? *The Wall Street Journal.*

Bernanke, B. S. et al. (2012). Opening remarks: Monetary policy since the onset of the crisis. In *Proceedings: Economic Policy Symposium Jackson Hole*, pages 1–22.

Bhattarai, S., Eggertsson, G. B., and Gafarov, B. (2015). Time consistency and the

duration of government debt: A signalling theory of quantitative easing. Technical report, National Bureau of Economic Research.

Bianchi, J. and Bigio, S. (2014). Banks, liquidity management and monetary policy. Technical report, National Bureau of Economic Research.

Brunnermeier, M. K. and Sannikov, Y. (2016). The I theory of money. Technical report, National Bureau of Economic Research.

Butler, A. W., Grullon, G., and Weston, J. P. (2006). Can managers successfully time the maturity structure of their debt issues? *The Journal of Finance*, 61(4):1731–1758.

Clouse, J., Henderson, D., Orphanides, A., Small, D. H., and Tinsley, P. A. (2003). Monetary policy when the nominal short-term interest rate is zero. *Topics in Macroeconomics*, 3(1).

Curdia, V. and Woodford, M. (2011). The central-bank balance sheet as an instrument of monetarypolicy. *Journal of Monetary Economics*, 58(1):54–79.

D'Amico, S., English, W., López-Salido, D., and Nelson, E. (2012). The Federal Reserve's large-scale asset purchase programmes: Rationale and effects. *The Economic Journal*, 122(564).

D'Amico, S. and King, T. B. (2013). Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply. *Journal of Financial Economics*, 108(2):425–448.

Di Maggio, M., Kermani, A., and Palmer, C. (2016). How quantitative easing works: Evidence on the refinancing channel. Technical report, National Bureau of Economic Research.

Drechsler, I., Savov, A., and Schnabl, P. (2014). A model of monetary policy and risk premia. Technical report, National Bureau of Economic Research.

ECB (2016a). Economic bulletin. February.

ECB (2016b). Economic bulletin. March.

Eggertsson, G. B. et al. (2003). Zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, 2003(1):139–233.

Falagiarda, M. and Reitz, S. (2015). Announcements of ECB unconventional programs: Implications for the sovereign spreads of stressed euro area countries. *Journal of International Money and Finance*, 53:276–295.

Faulkender, M. (2005). Hedging or market timing? Selecting the interest rate exposure of corporate debt. *The Journal of Finance*, 60(2):931–962.

Faulkender, M. and Petersen, M. A. (2006). Does the source of capital affect capital structure? *Review of Financial Studies*, 19(1):45–79.

Fratzscher, M., Duca, M. L., and Straub, R. (2016). ECB unconventional monetary policy: Market impact and international spillovers. *IMF Economic Review*, 64(1):36–74.

Gagnon, J., Raskin, M., Remache, J., Sack, B., et al. (2011). The financial market effects of the Federal Reserve's large-scale asset purchases. *International Journal of Central Banking*, 7(1):3–43.

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34.

Gertler, M., Kiyotaki, N., et al. (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of Monetary Economics*, 3(3):547–599.

Greenwood, R., Hanson, S., and Stein, J. C. (2010). A gap-filling theory of corporate debt maturity choice. *The Journal of Finance*, 65(3):993–1028.

Greenwood, R., Hanson, S. G., and Stein, J. C. (2016). The federal reserve's balance sheet as a financial-stability tool. In *Designing Resilient Monetary Policy Frameworks for the Future," Jackson Hole Symposium: Federal Reserve Bank of Kansas City*.

Greenwood, R. and Vayanos, D. (2014). Bond supply and excess bond returns. *The Review of Financial Studies*, 27(3):663–713.

Hamilton, J. D. and Wu, J. C. (2012). The effectiveness of alternative monetary policy tools in a zero lower bound environment. *Journal of Money, Credit and Banking*, 44(s1):3–46.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *The American Economic Review*, 103(2):732–770.

Hull, J., Predescu, M., and White, A. (2004). The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking & Finance*, 28(11):2789–2811.

Jiménez, G., Mian, A., Peydró, J.-L., and Saurina, J. (2014). The real effects of the bank lending channel.

Joyce, M., Lasaosa, A., Stevens, I., and Tong, M. (2011). The financial market impact of quantitative easing in the United Kingdom. *International Journal of Central Banking*, 7(3):113–161.

Khwaja, A. I. and Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *The American Economic Review*, 98(4):1413–1442.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. Technical report, National Bureau of Economic Research.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2012). The aggregate demand for treasury debt. *Journal of Political Economy*, 120(2):233–267.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2013). The ins and outs of lsaps.

Lam, W. R. (2011). Bank of Japan's monetary easing measures: Are they powerful and comprehensive? Technical report, International Monetary Fund.

Lenel, M. (2017). Safe assets, collateralized lending and monetary policy.

Platt, E. (2016, February 12). US and European junk bonds tumble. *Financial Times*.

Smith, R. (2016, February 5). Fallen angels loom over European bond market. *Reuters*.

Stein, J. C. (2012). Monetary policy as financial stability regulation. *The Quarterly Journal of Economics*, 127(1):57–95.

Swanson, E. T. (2011). Let's twist again: a high-frequency event-study analysis of operation twist and its implications for QE2. *Brookings Papers on Economic Activity*, 2011(1):151–188.

Swanson, E. T. (2015). Measuring the effects of unconventional monetary policy on asset prices. Technical report, National Bureau of Economic Research.

Szczerbowicz, U. et al. (2015). The ECB unconventional monetary policies: Have they lowered market borrowing costs for banks and governments? *International Journal of Central Banking*, 11(4):91–127.

Tang, T. T. (2009). Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement. *Journal of Financial Economics*, 93(2):325–351.

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of Money, Credit and Banking*, 1(1):15–29.

Ueda, K. (2012). The effectiveness of non-traditional monetary policy measures: The case of the Bank of Japan. *Japanese Economic Review*, 63(1):1–22.

Vayanos, D. and Vila, J.-L. (2009). A preferred-habitat model of the term structure of interest rates. Technical report, National Bureau of Economic Research.

Wallace, N. (1981). A Modigliani-Miller theorem for open-market operations. *The American Economic Review*, 71(3):267–274.

Whittall, C. (2016, January 18). A hint of trouble in European debt. *The Wall Street Journal*.

Woodford, M. (2016). Quantitative easing and financial stability. Technical report, National Bureau of Economic Research.

CHAPTER 3

FINANCIAL INTERMEDIATION AND FLIGHTS TO SAFETY

3.1 Introduction

Many existing contributions have interpreted countercyclical demand for riskless assets as reflecting time-varying costs of external financing. However, evidence from the behavior of the banking system exposes the empirical gaps of this theory. Indeed, banks can currently rely on virtually unlimited funds from central banks, but they are nevertheless holding record-high levels of excess reserves.

In this paper, I show that episodes of sudden portfolio re-allocation towards riskless securities are a robust feature of a game of incomplete information between investors and an intermediary. Specifically, I model a continuous time interaction between a population of investors and a large intermediary. Investors can purchase riskless securities and let the intermediary manage the remaining part of their capital. In exchange of its services, the intermediary receives a proportional fee, which is determined endogenously. Besides the risk-free asset, the intermediary has also access to a risky investment opportunity whose expected return (above the risk-free rate or below the risk-free rate) is its private information and is associated with the intermediary's type (good or bad). Investors do not observe the type of the intermediary or its portfolio allocation, but can observe a signal that coincides with the returns generated by the intermediary.

The key state variable of the model is the intermediary's reputation, i.e. in-

vestors' posterior beliefs that the intermediary is of the good type. The law of motion of reputation is endogenous and depends on the entire history of the intermediary's performance. This provides incentives for the intermediary to choose a portfolio allocation between riskless and risky assets in order to manipulate the distribution of the signal and, consequently, the evolution of its reputation. Clearly, in equilibrium, investors will rationally anticipate the strategy of the intermediary, neutralizing signal manipulation and giving rise to instances of signal jamming of the type discussed in Stein (1989).

The model generates episodes of hoarding of riskless assets that bear a qualitative similarity to those observed during financial crises (see Acharya and Merrouche, 2012 Ashcraft et al., 2011 and Beber et al., 2009 for empirical investigations of liquidity hoarding during the latest crisis). Hoarding, in my model, takes place in the asset side of the intermediary's balance sheet, as soon as it heavily invests in riskless securities in order to control the evolution of its reputation.

Contrary to existing literature, liquidity hoarding, in my model, is not directly associated with changes in exogenous state variables. Indeed, demand for riskless assets, in a Markovian equilibrium, negatively depends on the intermediary's reputation, which evolves endogenously as investors learn by observing returns. Therefore, hoarding will typically happen after a sequence of particularly low returns.

The equilibrium of the model is characterized by investors learning the type of the intermediary and by intermediaries of different types choosing the same portfolio allocation. Investors use Bayes' rule to update their beliefs about the type of the intermediary after observing the entire history of returns. As long as investors,

223

who are risk-neutral, expect the intermediary to deliver returns above the risk-free rate, the intermediary will survive and manage investors' capital. However, when beliefs fall below a threshold, the intermediary will be terminated, since investors will prefer to hold riskless securities. The bad intermediary, therefore, imitates the portfolio allocation of the good one. Indeed, if this was not the case, the two types would generate signals with different volatility and, hence, they could be immediately separated, leading to the termination of the bad type.

I provide analytical results showing that, when reputation is low enough and the risk of early termination is high, the intermediary, rather than gambling for resurrection, will become risk averse and minimize the risk in its portfolio. This outcome is in stark contrast with most of the existing literature on risk-shifting and is linked to the learning dynamics and the infinite horizon of the model[1]. Indeed, near the threshold below which the intermediary is terminated, the expected excess return on the intermediated asset is approximately zero, according to investors' beliefs. Therefore, a small amount of risk is sufficient for the good intermediary to exceed, on average, investors' expectations and, thus, to ensure that beliefs acquire a positive drift. Moreover, if a negative shock to returns, and hence to reputation, is realized, the good intermediary will be terminated and it will lose the option to signal its type through returns. Therefore, the optimal choice is to minimize the amount of portfolio risk.

In order to highlight the key mechanisms behind the model's results and test their robustness, I develop some variations and extensions of the basic set-up. I show that

---

1. Panageas and Westerfield (2009) have shown that the risk-taking incentives of convex payoff schemes may disappear if an asset manager has an infinite investment horizon.

lack of commitment about the intermediary's trading strategy is crucial in delivering the results of the paper. Indeed, whenever the intermediary can ex-ante commit to a portfolio allocation, no hoarding happens, since the signal-manipulation incentives disappear. Moreover, I show that hoarding can be observed also in those situations in which the intermediary cannot control the law of motion of investors' capital. This remarks that the relevant incentives entirely come from the informational asymmetries in the model.

To account for the recurring nature of liquidity hoarding episodes, I also develop a model where the type of the intermediary evolves stochastically. Indeed, if the type is fixed, beliefs will converge with probability one to either zero or one, reflecting the fact that beliefs must be asymptotically correct. This would imply that, if we consider a partial equilibrium model with a single intermediary, liquidity hoarding episodes are transient and will disappear in the long run. By allowing the intermediary's type to be time-varying, episodes of liquidity hoarding may appear with a cyclical pattern. My analysis suggests that the model is robust to this extension.

The rest of the paper is organized as follows. In section 3.2, I compare this paper to the existing literature. I then introduce the model in section 3.3, where I also give the definition of sequential equilibrium and provide a characterization. In section 3.4, I characterize and solve for a Markovian equilibrium where the state variables are the capital of investors and the reputation of the intermediary. Section 3.5 contains variations and extensions of the model. Finally, section 3.6 concludes. All proofs are collected in Appendix 3.A, while Appendix 3.B contains additional plots.

3.2 Related Literature

The relation between investors' confidence, reactions to news and liquidity hoarding has always been one of the reasons provided to explain historical patterns in demand for liquidity and riskless assets. Friedman and Schwartz (1971) wrote the following about the Great Depression:

> "Excess reserves, which in January 1931 had for the first time since 1929, when data became available, reached the $100 million level and had then declined as confidence was restored, again rose, reaching a level of $125-$130 million in June and July. Once bitten, twice shy, both depositors and bankers were bound to react more vigorously to any new eruption of bank failures or bank difficulties than they did in the final months of 1930."

A sharp rise in excess reserves has also been observed during the recent financial crisis (see Figure 3.7 in Appendix 3.B).

My model is able to rationalize these facts by appealing to one simple friction, i.e. the information asymmetry about the type of the intermediary and its portfolio allocation. This is in contrast with the banking and macro-finance literature, which usually appeals to moral hazard or unobservable cash flows as key frictions in financial markets (Bernanke and Gertler, 1989; Bernanke et al., 1999; Gertler and Karadi, 2011; Gertler and Kiyotaki, 2010).

A similar comparison can be made with the literature on intermediary asset pricing with frictions. The main result of this literature is the pro-cyclical variation

of demand for risky assets and the key driving force is the presence of leverage constraints or market incompleteness (Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2012, 2013). In my model, the driving forces behind the same kind of result are simply the inter-temporal incentives of intermediaries to manipulate the distribution of returns in bad times in order to avoid risks of reputation losses.

Flight to quality and liquidity hoarding have been subject to discussion under different perspectives in previous theoretical literature. In Bernanke et al. (1996) the key driver of these phenomena are time-varying agency costs. In Vayanos (2004) it is stochastic volatility coupled with risk of early liquidation of the intermediary. For Caballero and Krishnamurthy (2008) the reason behind liquidity hoarding lies in exogenous liquidity shocks and knightian uncertainty. Acharya et al. (2012), Acharya et al. (2007) and Acharya and Skeie (2011) focus instead on the role of external financing frictions and the variation in credit risk, investment opportunities and roll-over risk, respectively, to explain variation in demand for safe assets. Finally, in Gale and Yorulmazer (2013) the absence of contingent markets for the provision of liquidity creates incentives for investors to occasionally hoard liquidity for precautionary or speculative reasons.

This paper is also closely related to the literature on reputation in dynamic games. Fudenberg and Levine (1992) provide the classic benchmark in a discrete time model, where they derive bounds for the equilibrium payoffs of the rational large player, as its discount rate tends to zero. Faingold and Sannikov (2011) extend this model to continuous time and obtain a sharper characterization of the equilibrium payoffs.

These models share the common feature that reputation does not carry a premium

227

per se, but it just affects the actions of players. Board and Meyer-ter Vehn (2013) model a reputation game between consumers and a firm whose product quality is private information. The firm enjoys reputational dividends, as reputation affects the market price of the product. Similarly, in my paper, reputation will directly affect the cash flow of the intermediary through the fee paid by investors.

My approach also bears connections with the recent literature of continuous time games. However, besides the fact that my set-up differs from the typical principal-agent problem, an additional point of departure is the fact that, in my model, the intermediary controls also the volatility of the signal, and not only its drift. Holmstrom and Milgrom (1987) is the first contribution to the literature of continuous time games where the signal follows a diffusion process whose probability distribution is controlled by an agent. Later, Sannikov (2007) and Sannikov (2008) extended this framework to more general settings and showed that the key state variable in this class of models is the continuation value of the agent. Indeed, this is the relevant variable to take into consideration in the provision of dynamic incentives. DeMarzo and Sannikov (2006) apply these techniques to the optimal design of securities of firms and He (2009) extends the analysis to firms growing as a geometric Brownian motion whose drift is controlled by the agent.

The type of economic interaction explored in this paper is very similar to the typical setting in the delegated asset management literature. Hugonnier and Kaniel (2010) is probably the closest, in spirit, to my paper. Indeed, they characterize the optimal portfolio allocation of an asset manager who has to dynamically raise funds from external investors. However, their implicit assumption is that the asset manager

228

is able to ex-ante commit to a dynamic trading strategy and that, in choosing it, the asset manager takes into account how the strategy will affect the dynamic incentives of investors. This leads to the possibility of sub-game imperfection. In my model, the equilibrium is subgame perfect and is characterized by imperfect information and by the impossibility of the asset manager to commit to a state contingent trading strategy.

Vayanos (2004) explores the time variation of liquidity premia in a general equilibrium model with intermediation. Stochastic volatility is the main state variable of the model and intermediaries are subject to early termination according to an exogenously specified rule, that is based only on the short-term performance of the intermediary. In my model, time-varying demand for safe assets does not depend on stochastic volatility, but only on the endogenous evolution of reputation. Furthermore, in my set-up, the termination rule is endogenous and depends on the entire history of the intermediary's performance.

Other papers (Basak and Cuoco, 1998; Cuoco and Kaniel, 2011; Kaniel and Kondor, 2013) have analyzed the asset pricing implications of delegated asset management. Despite I limit the analysis to a partial equilibrium environment, I provide a micro-foundation to the asset manager's compensation when he faces a set of risk neutral and competitive investors. This makes the fees paid to the asset manager naturally dependent on the historical performance of the funds, while in previous contributions fees where either fixed or assumed to depend on past performance through an exogenously specified function. Fees, in my model, represent a premium that the intermediary receives for its reputation.

229

Indeed, contributions in the active portfolio management literature, e.g. Berk and Green (2004), showed that the presence of competitive capital markets is able to dramatically change the empirical asset pricing implications of models with intermediation. While, in Berk and Green (2004), equilibrium is achieved through changes in the size of the fund and decreasing returns to scale, in my model adjustment happens through the price, i.e. the fee, paid by investors to access the fund's services.

3.3 Model

I consider a partial equilibrium model where a large intermediary and a population of investors are present. Investors have the choice of either investing in a riskless asset or let an intermediary manage their capital. The intermediary manages investors' fund in exchange of a proportional fee that, in equilibrium, constitutes a premium for its reputation. The intermediary can be one of two types and investors cannot observe the type of the intermediary.

3.3.1 Investment Opportunities

$$\frac{dS_t^0}{S_t^0} = rdt, \quad r > 0,$$

for a given S_0^0.

The other two assets are risky. The price of the first risky asset, defined as the good asset, evolves according to

$$\frac{dS_t^1}{S_t^1} = \mu_1 dt + \sigma dW_t,$$

for a given S_0^1, while the second asset, defined as the bad asset, follows

$$\frac{dS_t^2}{S_t^2} = \mu_2 dt + \sigma dW_t,$$

for a given S_0^2. Note that both risky assets have the same volatility and are exposed to the same Weiner process.

The following assumption provides the reason why the two assets are defined as good and bad.

Assumption 3.1. *In an economy where no individual is risk-lover, the bad asset is dominated by the other two, i.e.*

$$\mu_1 > r > \mu_2.$$

Investors can only hold positive amounts of the riskless asset and let an intermediary manage the remaining part of their portfolio. The intermediary can be one of two types, good or bad. Both types can invest in the riskless security, but they differ in terms of the risky investment opportunity that is available to them. Indeed, the good intermediary can invest in the good asset but not in the bad one, while the bad intermediary can invest in the bad asset, but not in the good one.

3.3.2 Players

The economy is populated by a unit measure of small and atomistic investors so that none of them, taken individually, is able to influence the incentives of the intermedi-

ary.

Investors are risk neutral and discount future utility at rate r, thus evaluating a consumption process $(c_t)_{t\geq 0}$, conditional on the history of the game up to time t, according to

$$\mathrm{E}\left[\int_0^\infty e^{-rs} c_s \, ds \,\Big|\, \mathcal{F}_t^I\right].$$

where $(\mathcal{F}_t^I)_{t\geq 0}$ is the investors' information filtration, which will be fully characterized in section 3.3.3.

Investors are endowed with a stock of capital K_0 at time zero. At each time t, they decide which fraction of their capital that will be invested in riskless assets. The remaining part will be managed by the intermediary from time t to time $t + dt$. Let $\theta_t \in [0, 1]$ be the fraction of capital that investors delegate to the intermediary and let dR_t be the return generated by the intermediated assets. Then, the law of motion of investors' capital is

$$\frac{dK_t}{K_t} = (1 - \theta_t)rdt + \theta_t(dR_t - f_t dt). \tag{3.1}$$

f_t is the proportional fee received by the intermediary for managing capital from t to $t + dt$. The fee in paid up-front at time t and investors will receive the full return, dR_t, generated by the intermediated assets. The fee will be endogenously determined by a break-even condition for investors. The existence of fees will provide incentives for the bad intermediary to keep operating despite the socially wasteful investment opportunity. The intermediary cannot save on its own account and immediately consumes all the fees it receives. Therefore fees have to be non-negative.

Once investors decide the amount of assets that the intermediary will manage, the latter will be free to choose a portfolio allocation between the riskless asset and the risky investment opportunity available. In this sense, the portfolio allocation of the intermediary is not contractible.

Indeed, the intermediary will choose a process for the fraction of risky assets in its portfolio, which I call portfolio risk and denote with $(i_t)_{t\geq0}$, after the investors' capital has been transferred to the intermediary and after fees have been collected. Investors cannot directly observe the choice of portfolio risk. Moreover, even if it was revealed, there would be no mechanism to impose a penalty on an intermediary that deviates from a previously agreed portfolio risk.

The intermediary is risk neutral and discounts future consumption at rate ρ. Given a process for the assets under management, $(\theta_t K_t)_{t\geq0}$, the intermediary receives a lifetime utility, conditional on the history of the game up to time t, of

$$\mathrm{E}\left[\int_0^\infty e^{-\rho s}\theta_s K_s f_s\,ds\Big|\mathcal{F}_t\right].\tag{3.2}$$

where $(\mathcal{F}_t)_{t\geq0}$ represents the intermediary's information filtration (more on this in section 3.3.3).

Therefore, given an investment strategy $(i_t)_{t\geq0}$ for the intermediary, the return on the assets under management is described by

$$dR_t = rdt + i_t(h\mu_1 + (1-h)\mu_2 - r)dt + i_t\sigma dW_t.\tag{3.3}$$

$h \in \{0,1\}$ is random variable, drawn at time zero, that defines the type of the

233

intermediary. $h = 1$ indicates that the intermediary is of the good type and, thus, has access to the good investment opportunity. $h = 0$ means that the intermediary is bad and can therefore invest in the bad asset.

Given the risk neutrality of the players, it is necessary that the portfolio risk, i_t, lies in a compact set in order to guarantee the existence of a solution to the portfolio allocation problem, at least in the limiting case of perfect information with a good intermediary. Hence, I assume the following.

Assumption 3.2. *The intermediary is subject to a minimum and a maximum portfolio risk constraints, i.e.*

$$i_t \in [\hat{i}, \bar{i}] \quad \forall t.$$

with $0 < \hat{i} < \bar{i}$.

A positive upper bound for i_t can be interpreted as a limit on risk-free borrowing. While a no-short selling constraint would impose $\hat{i} = 0$, this model actually requires the slightly stronger assumption that $\hat{i}$ is strictly positive. The motivation is related to the learning part of the model that will be described in section 3.3.4 and in Proposition 3.2 in particular. To explain briefly, the fact that the intermediary has always to hold some risk in its portfolio guarantees that returns are always stochastic and investors will always learn from them. This avoids theoretical complications in the dynamic game when considering deviation to a portfolio risk $i_t' > 0$ when the equilibrium risk is $i_t = 0$. In these cases, the equilibrium returns are deterministic and investors do not learn from them in equilibrium. If, however, the intermediary deviates, then it will provide investors with a return that, with probability 1, differs

234

from the equilibrium (deterministic) one. Strong assumptions would then be needed to pin down the off-equilibrium beliefs of investors.

A further assumption will be needed to ensure the existence of a solution to the model in the perfect information case. As it will become clear in section 3.4.1, when the intermediary is of the good type and there is perfect information, in equilibrium fees will be constant and capital will grow as a geometric Brownian motion with drift r. The following assumption is then necessary for the continuation value of the good intermediary to be bounded under perfect information.

Assumption 3.3. *The intermediary is more impatient than investors, i.e.*

$$\rho > r.$$

In the next subsection, I will precisely describe the information structure of the game and formalize the definition of the strategies of the players.

3.3.3 Information Structure

There are two sources of asymmetric information in this game. The first one is that the type of the intermediary (good or bad) is private information of the intermediary itself. The second one is that investors do not observe the portfolio allocation i_t of the intermediary, but only the total return dR_t. This implies that investors' strategies cannot directly depend on the type of the intermediary they are facing, or on the history of the Wiener process $(W_t)_{t \geq 0}$.

To set the notation, let $(\Omega, \mathcal{F}^*, P)$ be a probability space and let the Wiener

process $(W_t)_{t\geq0}$ and the type of the intermediary h be independent random variables on $(\Omega, \mathcal{F}^*)$.

In order to formally define the strategies of the investors, let $(\mathcal{F}_t^I)_{t\geq0}$ be the filtration generated by the process $(R_t)_{t\geq0}$, possibly augmented by the collection of P-null sets. An action θ_t for the investors at time t is an $\mathcal{F}_t^I$-measurable function that maps histories of returns up to time t into the action space of delegated portfolio share, $[0, 1]$.

The intermediary is instead endowed with a larger filtration, namely a filtration $(\mathcal{F}_t)_{t\geq0}$ that is generated by h and $(W_t)_{t\geq0}$ and is possibly augmented by the collection of P-null sets. Therefore, an action i_t for the intermediary at time t is an $\mathcal{F}_t$-measurable function that maps its type and histories of the Wiener process up to time t into the action space $[\hat{i}, \bar{i}]$. Let i_t^G and i_t^B indicate the action i_t conditional on the type being good ($h = 1$) and bad ($h = 0$), respectively. Since I consider only pure strategy equilibria, the filtration $(\mathcal{F}_t)_{t\geq0}$ coincides with the filtration generated by h and $(R_t)_{t\geq0}$[2] and, hence, it is without loss of generality to assume that the actions of the two types, i_t^G and i_t^B, are public, i.e. they depend on the history of the signal $(R_s)_{0\leq s\leq t}$.

Finally, the fee f_t is defined as an $\mathcal{F}_t^I$-measurable random variable that represents the premium investors are willing to pay to the intermediary to have access to its intermediation services.

Investors do not know the type of the intermediary they are facing. At time zero, they start with a prior $\phi_0 \in [0, 1]$ that the intermediary is of the good type.

---

2. See Appendix 3.A.1 for a proof.

Conditional on the intermediary having a positive amount of assets under management, investors will receive a signal, coinciding with the return generated by the intermediary, dR_t. On the basis of the signal and of the equilibrium strategies of the two types, investors will update their beliefs. A belief process $(\phi_t)_{t \geq 0}$ is therefore a stochastic process, adapted to $(\mathcal{F}_t^I)_{t \geq 0}$, representing the probability that investors assign, at each time t, to the good intermediary. I refer to ϕ_t as the reputation of the intermediary at time t.

3.3.4 Sequential Equilibrium Definition and Characterization

I will now define and characterize a sequential equilibrium of this game. The following definition implicitly assumes the existence of market clearing mechanism for the allocation of investors' funds. Moreover, it assumes that, whenever investors are indifferent between the riskless asset and the intermediary, they will let the intermediary manage their whole portfolio, provided that the intermediary generates a strictly positive expected excess return, according to their information[3].

Definition 3.1 (Public Sequential Equilibrium). *A public sequential equilibrium consists in a fee process $(f_t)_{t \geq 0}$, a process for the fraction of capital managed by the intermediary $(\theta_t)_{t \geq 0}$, a portfolio risk process $(i_t)_{t \geq 0}$ and a belief process $(\phi_t)_{t \geq 0}$, such that, for all times $t \geq 0$ and after every history of the game up to time t, the following conditions hold.*

---

3. In a partial equilibrium setting with risk-neutral investors like this one, it is not possible to precisely pin down quantities. One way to justify this assumption, is to think that the intermediary can always request an infinitesimally smaller fee and break the tie in its favor. This is the reason why, according to Definition 3.1, in equilibrium $\theta_t = 1$ if and only if $f_t > 0$.

(i) Investors break even,

$$f_t = \mathrm{E}[i_t(h\mu_1 + (1-h)\mu_2 - r)|\mathcal{F}_t^I].$$

(ii) $\theta_t = 1$ if and only if $\mathrm{E}[i_t(h\mu_1 + (1-h)\mu_2 - r)|\mathcal{F}_t^I] > 0$, $\theta_t = 0$ otherwise.

(iii) Given the public strategy profile $(i_t)_{t\geq0}$ and an initial $\phi_0 \in [0, 1]$, beliefs $(\phi_t)_{t\geq0}$ are updated using Bayes' rule and are consistent with the public strategy profile,

$$\phi_t = \mathrm{E}[h|\mathcal{F}_t^I].$$

(iv) $(i_t)_{t\geq0}$ maximizes the intermediary's lifetime utility (3.2) given (i), (ii),(iii) and the law of motion of capital (3.1).

We have seen that we can restrict our attention to equilibria in public strategies. This means that the strategies of the two types may differ, but they will depend on the history of the process R_t. The following proposition provides a stronger characterization of the equilibrium strategies of the two types.

Proposition 3.1. *In equilibrium, i_t is $\mathcal{F}_t^I$ measurable P-a.s. and, therefore, $i_t^G = i_t^B$ P-a.s..*

The intuition for this result is that, being $(i_t)_{t\geq0}$ related to the volatility of the process $(R_t)_{t\geq0}$, an observer is able to learn about it very quickly thanks to the high frequency movements of $(R_t)_{t\geq0}$. Therefore, if the two types choose different portfolio risks, consistency of beliefs will impose to assign a correct and degenerate

posterior probability distribution over types. Since the bad intermediary destroys value for investors, the bad intermediary will be terminated if it is separated. Given this belief formation rule, the bad type has incentives to pool with the good one.

Since zero measure deviations do not change the lifetime utility of the players, I will simply assume that the two types pool at every time t. We can therefore simply denote as $(i_t)_{t \geq 0}$ the common choice of portfolio risk process of the two types.

Proposition 3.1, however, leaves a large number of possible equilibria. Indeed, given investors' expectations about the process $(i_t)_{t \geq 0}$ chosen by the good type, both types will have an incentive to pool to such process in order not to be considered bad and terminated by investors.

For the purposes of this paper, I will simply assume that the bad intermediary imitates the choice of the good one. The good intermediary is free to choose the process $(i_t)_{t \geq 0}$ that maximizes its lifetime utility, without concerns of seeing its own reputation decreased or increased because of its choice of $(i_t)_{t \geq 0}$. In other words, the good intermediary chooses its portfolio risk while assuming that investors update their beliefs exclusively on the basis of the returns they observe. Investors understand this and understand that the bad intermediary will imitate the good one. Therefore, they will not penalize deviations towards strategies that are optimal for the good intermediary, given the way in which investors learn from returns. I leave the investigation of more general equilibrium refinement concepts for future work[4].

Once the intermediary chooses its portfolio allocation, it will generate a return

---

4. Extending the divinity refinement concept of Banks and Sobel (1987) to continuous time games seems to be a promising way to pin down the type of pooling equilibrium that I assume in this paper.

that will be delivered to investors without any agency friction. Since different types of intermediaries induce different probability distribution over returns, investors will exploit the history of returns to learn the type of the intermediary. Standard filtering results, extensively used in the literature on learning and continuous time strategic experimentation (Bolton and Harris, 1999; Hansen and Sargent, 2011; Pástor and Veronesi, 2009; Veronesi, 1999), lead to the following Proposition.

Proposition 3.2. *Given a prior at time 0 that the intermediary is good $\phi_0 \in [0, 1]$, and given an equilibrium strategy profile $(i_t)_{t \geq 0}$, a belief process $(\phi_t)_{t \geq 0}$ is consistent with the strategy profile if, when $\theta_t > 0$, it satisfies the stochastic differential equation*

$$d\phi_t = \varphi(\phi_t) \, (i_t\sigma)^{-1} \, (dR_t - \mu(i_t, \phi_t)dt) \,, \tag{3.4}$$

with

$$\varphi(\phi_t) = \phi_t(1 - \phi_t)\sigma^{-1}(\mu_1 - \mu_2)$$

and

$$\mu(i_t, \phi_t) = r + i_t \left[\phi_t\mu_1 + (1 - \phi_t)\mu_2 - r\right]$$

and where the expression for dR_t is given by (3.3), while, when $\theta = 0$, $d\phi_t = 0$.

The terms $\varphi(i_t, \phi_t), (i_t\sigma)^{-1}$ and $\mu(i_t, \phi_t)$, as well as, the distribution of dR_t depend on i_t. However, they do so for different reasons. Indeed, the quantities $\varphi(i_t, \phi_t), (i_t\sigma)^{-1}$ and $\mu(i_t, \phi_t)$ are set by investors as a function of the strategy played in equilibrium by the intermediary. It follows that those terms will be taken as given by the intermediary when it has to choose its trading strategy. On the contrary, the

distribution of dR_t directly depends on the action of the intermediary. The intermediary, therefore, will consider how the distribution of dR_t changes when it considers deviations from the equilibrium strategies.

Of course, an equilibrium must be such that the intermediary has no incentives to deviate. But, by keeping the distinction clear, I intend to stress the main channel through which signal manipulation incentives arise.

Suppose that the intermediary is contemplating a deviation $(i'_t)_{t \geq 0}$ from the equilibrium $(i_t)_{t \geq 0}$. This will induce a different probability distribution for the signal dR_t so that, under the intermediary's information filtration, the law of motion of beliefs is given by

$$d\phi_t = \varphi(\phi_t)\,(i_t\sigma)^{-1}\,[r + i'_t(h\mu_1 + (1-h)\mu_2 - r) - \mu(i_t, \phi_t)]dt + \varphi(\phi_t)\,(i_t\sigma)^{-1}\,(i'_t\sigma)dW_t.$$

In equilibrium, the intermediary must have no incentive to choose a trading strategy that is different from the equilibrium one. Therefore, under the information filtration of the intermediary, the equilibrium law of motion of beliefs is given by

$$d\phi_t = \varphi(\phi_t)(h - \phi_t)\frac{\mu_1 - \mu_2}{\sigma}dt + \varphi(\phi_t)dW_t.$$

The drift in the latest equation is positive if the intermediary is good, i.e. if $h = 1$, while it is negative when the intermediary is bad, i.e. when $h = 0$. This reflects the fact that, while under the information filtration of investors beliefs are martingales, under the information filtration of the intermediary they converge a.s. to either 0 or 1, depending on whether the intermediary is bad or good, respectively.

241

While the bad intermediary has to follow the trading strategy chosen by the good one, the latter can choose a trading strategy in order to control the evolution of capital and beliefs by manipulating the distribution of returns dR_t. Indeed, by controlling the evolution of K_t and ϕ_t, the intermediary is controlling the process for its cash flow $K_t f_t$. For example, the good intermediary may have an incentive to choose a low i_t in order to reduce the volatility of the signal and, consequently, of beliefs and capital, despite this will reduce their drift. However, in equilibrium, these incentives will be understood by investors who will then scale the signal by the appropriate liquidity ratio i_t. This could give rise to a phenomenon of signal-jamming in the demand for riskless assets. The intermediary may try to manipulate the distribution of the signal by demanding less risk, but such manipulation will be offset in equilibrium by the rational expectations of the investors.

The discussion so far has been quite general and players' actions at time t may depend on the entire history of returns up time t. In the next section, I discuss the properties of an equilibrium that is Markovian in capital and beliefs.

3.4   Markovian Equilibrium

Investors' capital and beliefs naturally embody the entire history of the game and constitute natural state variables of the model. In an equilibrium that is Markovian in these state variables, the intermediary has to consider two elements which affect the amount of assets under management and the fees in future periods: the growth rate of investors' capital (3.1) and the change in its reputation (3.4).

Thanks to the functional form of the intermediary's utility, it is possible to guess and verify[5], that the continuation value of the intermediary will be linear in capital,

$$V(K, \phi) = K g(\phi), \tag{3.5}$$

and that the players' strategies will be functions of the intermediary's reputation only and not of the level of capital, since capital is essentially a scaling variable. Therefore, let $i(\phi_t)$ denote the value of the intermediary's risk at time t in a Markovian equilibrium and let $\theta(\phi_t)$ be the value of θ_t in a Markovian equilibrium.

The break even condition (i) in Definition 3.1, will therefore become

$$f_t = f(\phi_t) \equiv i(\phi_t) \left[\phi_t \mu_1 + (1 - \phi_t)\mu_2 - r \right]. \tag{3.6}$$

Since $i(\phi) > 0$ for all ϕ by Assumption 3.2, then condition (ii) of Definition 3.1 can be expressed in terms of a threshold $\bar{\phi}$ for the intermediary's reputation below which the expected excess return of the intermediary's portfolio, given investors' information, is negative for any possible value of $i(\phi)$. Such threshold is given by

$$\bar{\phi} = \frac{r - \mu_2}{\mu_1 - \mu_2}. \tag{3.7}$$

For any $\phi > \bar{\phi}$ the intermediary will generate, according to investors' information, a strictly positive expected excess return. By condition (i) in Definition 3.1, this excess return equals $f(\phi)$ and represents reputation rent of the intermediary, collected

---

5. See Proposition 3.3 below.

in the form of fees, while investors will break even. If instead $\phi \leq \bar{\phi}$, investors expect the intermediary to generate negative excess returns. Since the intermediary cannot pay investors to allow it to experiment, investors will set $\theta(\phi) = 0$, beliefs will be no longer updated and the intermediary will be terminated forever.

In the rest of this section, I will derive boundary conditions for the limiting cases of $\phi = \bar{\phi}$ and $\phi = 1$. I will then characterize the equilibrium in the open set $(\bar{\phi}, 1)$ and provide numerical results. Since I am assuming that the bad intermediary pools on the trading strategy that is optimal for the good intermediary given (3.1) and (3.4), I can solve for an equilibrium by simply focusing on the trading strategy and continuation value of the good intermediary.

3.4.1 Boundary Conditions and Degenerate Beliefs Case

It is straightforward to derive the boundary condition at $\phi = \bar{\phi}$. Since, at that point, the intermediary is terminated it must be the case that

$$g(\bar{\phi}) = 0 \tag{3.8}$$

and that $g(\phi) = 0$ for all $\phi \in [0, \bar{\phi}]$.

I now derive the equilibrium continuation value and trading strategy in the perfect information set-up. On the one hand, this provides a benchmark towards which we can compare the imperfect information equilibrium outcomes. On the other hand it provides a second boundary condition that is needed to solve the model with imperfect information.

Let $\phi = 1$. By equation (3.4), this implies that there will be no learning. Given the guess (3.5), the Hamilton-Jacobi-Bellman (HJB) equation is

$$\rho g(1) = \max_{i \in [\hat{i}, \bar{i}]} \{f(1) + g(1)(r + i[(\mu_1 - r) - f(1))]\}.$$

It is clear that the solution requires $i(1) = \bar{i}$ and the continuation value exists under Assumption 3.3 with

$$g(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r}. \tag{3.9}$$

3.4.2 Markovian Equilibrium Characterization

I will now characterize a Markovian equilibrium of the game. I will show that, provided that a function $g(\phi)$ exists and that it satisfies certain conditions, then the continuation value is indeed given by $Kg(\phi)$. Finally, I will provide analytical results about the shape of $g(\phi)$ and the equilibrium portfolio risk $i(\phi)$ for ϕ in a neighborhood of $\bar{\phi}$. Contrary to what standard model with convex incentive schemes predict, the intermediary displays risk aversion near the termination threshold and, in equilibrium, it will keep as little risk as possible in its portfolio.

Since the intermediary maximizes (3.2) subject to (3.1) and (3.4), the HJB equa-

245

tion is

$$\rho g(\phi) = \max_{i \in [\hat{i}, \bar{i}]} \left\{ \underbrace{f(\phi)}_{\text{flow payoff}} + \right.$$

$$+ \underbrace{g'(\phi)\varphi(\phi)\frac{[i(\mu_1 - r) - f(\phi)]}{i(\phi)\sigma}}_{\text{first order effect of } d\phi} +$$

$$+ \underbrace{g(\phi)(r + i(\mu_1 - r) - f(\phi))}_{\text{first order effect of } dK} + \qquad (3.10)$$

$$+ \underbrace{\frac{1}{2}g''(\phi)\varphi(\phi)^2\frac{i2}{i(\phi)^2}}_{\text{second order effect of } d\phi} +$$

$$\left. + \underbrace{g'(\phi)\varphi(\phi)\sigma\frac{i2}{i(\phi)}}_{\text{interaction effect of } d\phi \text{ and } dK} \right\}.$$

in the interval $(\bar{\phi}, 1)$. It is important to bear in mind the difference between $i(\phi)$ and i in equation (3.10). $i(\phi)$ represents the equilibrium trading strategy with respect to which beliefs are consistent. It is therefore taken as given by the intermediary. i is the choice variable of the intermediary, which is chosen in order to control the distribution of returns and, consequently, the distribution of capital growth and beliefs changes. Of course, in equilibrium, the optimal choice of i by the intermediary must coincide with $i(\phi)$.

An equilibrium $i(\phi)$ is therefore defined as a fixed point

$$i(\phi) \in \arg\max_{i \in [\hat{i}, \bar{i}]} \left\{ g'(\phi)\varphi(\phi)\frac{i(\mu_1 - r)}{i(\phi)\sigma} + g(\phi)i(\mu_1 - r) \right.$$

$$\left. + \frac{1}{2}g''(\phi)\varphi(\phi)^2\frac{i2}{i(\phi)^2} + g'(\phi)\varphi(\phi)\sigma\frac{i2}{i(\phi)} \right\} \quad (3.11)$$

with $f(\phi) = i(\phi)[\phi\mu_1 + (1-\phi)\mu_2 - r]$.

Consequently, in equilibrium g must solve the following second order differential equation

$$
\begin{aligned}
\rho g(\phi) = {} & f(\phi) + \\
& + g'(\phi)\phi(1-\phi)^2 \left(\frac{\mu_1 - \mu_2}{\sigma}\right)^2 + \\
& + g(\phi)[r + i(\phi)(1-\phi)(\mu_1 - \mu_2)] + \\
& + \frac{1}{2}g''(\phi)\phi^2(1-\phi)^2 \left(\frac{\mu_1 - \mu_2}{\sigma}\right)^2 + \\
& + g'(\phi)\phi(1-\phi)(\mu_1 - \mu_2)i(\phi)
\end{aligned}
\tag{3.12}
$$

where $i(\phi)$ is defined as in (3.11) and $f(\phi) = i(\phi)[\phi\mu_1 + (1-\phi)\mu_2 - r]$.

For the purposes of this paper, I will work under the assumption that there exists a solution to (3.12). A careful exploration of questions related to existence and uniqueness of the solution to the HJB equation will be subject to future research.

Assumption 3.4. *There exists a bounded function* $g\colon [\bar{\phi}, 1] \to \mathbb{R}$ *that is twice continuously differentiable in* $(\bar{\phi}, 1)$ *and that satisfies* (3.12) *with boundary conditions* $g(\bar{\phi}) = 0$ *and* $g(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r}$.

So far, I have guessed that the continuation value of the good intermediary is linear in capital and heuristically proceeded to characterize an equilibrium by means of equations (3.11) and (3.12). The following proposition provides a verification theorem that confirms that my guess and characterization are valid if Assumption 3.4 holds.

Proposition 3.3. *Under Assumption 3.4, if* $(\phi_t)_{t\geq 0}$ *evolves according to* (3.4) *with*

$i_t = i(\phi_t)$, where $i(\phi)$ solves (3.11) for any $\phi \in (\bar{\phi}, 1)$, then $V(K_t, \phi_t) = K_t g(\phi_t)$ is the continuation value of the good intermediary at time t and $(i(\phi_t))_{t \geq 0}$ is the equilibrium trading strategy.

A second order differential equation like (3.12) does not lend itself to straightforward analytical solutions. However, it is possible to characterize the shape of $g(\phi)$ when ϕ is sufficiently close to the termination threshold $\bar{\phi}$. Similarly, is it possible to solve for the equilibrium portfolio risk $i(\phi)$ in the proximity of $\bar{\phi}$.

While someone may expect the good intermediary to have incentives to risk-shift and gamble as much as possible in order to leave the neighborhood of $\bar{\phi}$, the following Proposition shows that exactly the opposite happens.

Proposition 3.4. *Under Assumption 3.4, there exists an $\epsilon > 0$ such that*

$$g''(\phi) < 0 \quad \forall \phi \in (\bar{\phi}, \bar{\phi} + \epsilon)$$

and

$$i(\phi) = \hat{i} \quad \forall \phi \in (\bar{\phi}, \bar{\phi} + \epsilon).$$

The proposition shows that the marginal value of capital $g(\phi)$ is concave near $\bar{\phi}$, so that the intermediary is averse to the risk of reputation losses in that region. Furthermore, rather than gambling for resurrection, the intermediary chooses the safest feasible portfolio near the termination threshold. This is in stark contrast with models of risk-shifting with convex payoffs.

The reason behind this result is that, when ϕ is close enough to $\bar{\phi}$, a very small portfolio risk is sufficient for the good intermediary to give a positive drift to investors'

248

beliefs. Moreover, the intermediary has no incentive to increase the volatility of the belief process above the equilibrium one. Indeed, if a negative shock is realized and the intermediary is terminated, the good intermediary would lose the option to signal its type and gain reputation in the future. Therefore, near the threshold, for any equilibrium $i(\phi)$, the unconstrained maximizer of the right-hand side of (3.11) is strictly smaller than $i(\phi)$ itself and, hence, the only possible equilibrium is the one where $i(\phi) = \hat{i}$.

The result of Proposition 3.4 is consistent with evidence of liquidity hoarding and flight to quality during financial crises. Given a sequence of bad returns that drive down intermediaries' reputation, the model predicts a shift of their portfolio towards safe assets, which, in most cases, consist in currency or short-term government bonds.

In section 3.5.1, I introduce a variation of the model where the intermediary can commit to a state contingent trading strategy. A comparison between Proposition 3.4 and Proposition 3.5 in section 3.5.1 highlights the role of incomplete contracts in generating hoarding of safe assets. Indeed, if the intermediary could commit to a trading strategy, in equilibrium the intermediary would always invest in risky assets up to the limit $\bar{i}$. This is because the signal manipulation incentives disappear and only concerns about the growth rate of capital remain. On the contrary, when the intermediary cannot commit, it will have incentives to trade off expected returns for lower volatility in order to control the evolution of beliefs.

3.4.3   Numerical Results

I will now report numerical solutions to ODE (3.12).

(a) Marginal Value of Capital   (b) Risky Assets to Total Capital
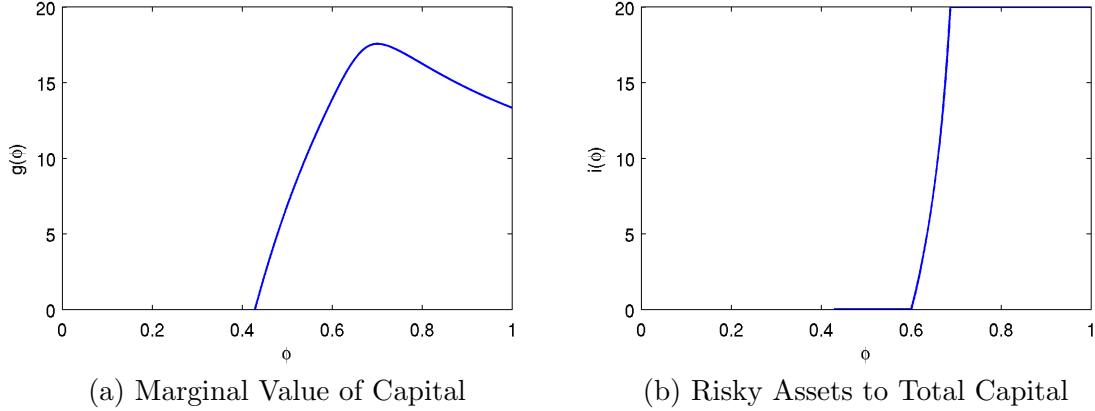
Figure 3.1: Case 1. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

Figure 3.1 shows the marginal value of capital $g(\cdot)$ of the good intermediary and the portfolio risk as a function of the beliefs of investors. The model in the Figure is parameterized as follows: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$. As a robustness check, in Appendix 3.B, I report numerical results for eight alternative parameterizations of the model. The shape of the value function and the shape of the demand for risky assets are robust to changes in the parameters.

We can immediately notice a pattern that we can call of liquidity hoarding. For high levels of reputation, the intermediary will invest in the risky asset as much as possible. However, there is a level of reputation below which the demand for risky assets decreases very rapidly until it reaches its lower bound.

It is easy to see how this mechanism is likely to play a role during financial crises. Indeed, the dynamics of the reputation are tightly linked to the history of returns

on the assets under management. Given a long enough sequence of bad return, the reputation of the intermediary may fall below a critical level and trigger a run towards liquid and safe assets.

For illustrative purposes, Figures 3.2 and 3.3 show the time series of the demand for risky assets in a simulated economy, when the returns are generated by the good and by the bad intermediary, respectively. This pattern, conditional on the bad type, is qualitatively similar to the evolution of excess reserves in Figure 3.7. There we see that, in a short interval of time, excess reserves saw a dramatic increase. This would be consistent with ϕ moving through the critical region from above, thus making i fall and demand for riskless assets increase.

Since beliefs are asymptotically correct[6], the "typical" time series, when the type is fixed and ϕ_0 is high enough, will be characterized by small probabilities of liquidity hoarding if the intermediary is good, and high probability of liquidity hoarding if the type is bad, but with no recurring patterns. One way to obtain recurring episodes of hoarding is to allow the type to be time-varying. Preliminary results from this model are available in section 3.5.4 and show the robustness of the model to this extension.

It is interesting to note that the function g is decreasing for high values of ϕ and this region coincides with the one where $i(\phi) = \bar{i}$. The reason is that, once i hits the upper limit, it can no longer be increased. Therefore, the expected return on the assets under management (under the probability measure induced by the good intermediary) stays constant but the fees raised by the intermediary keep increasing

---

6. In this model, learning stops when $\phi = \bar{\phi}$, so, conditional on the bad type, beliefs converges with probability one to $\bar{\phi}$.

(a) Beliefs

(b) Risky Assets to Total Capital
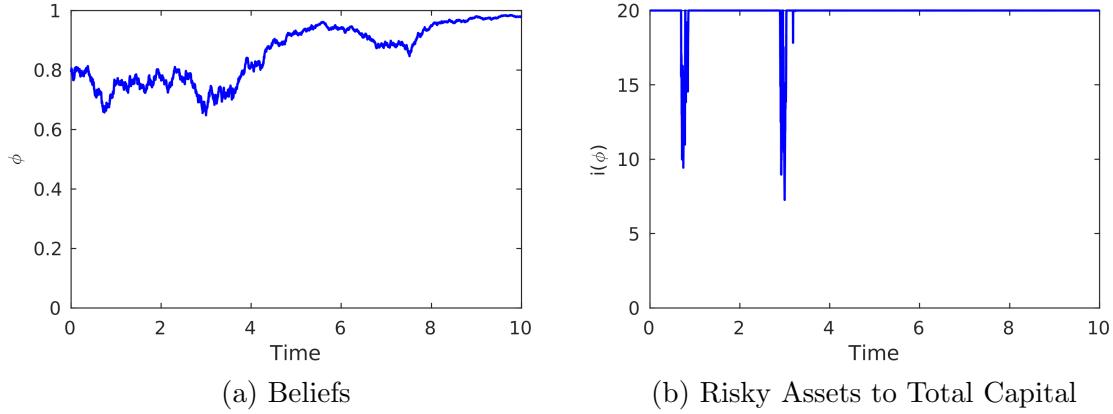
Figure 3.2: Good Intermediary. Time series of investors's beliefs and demand for risky assets as share of total capital under management when returns are generated by the good type. Parameter values are : $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

as ϕ increases. This has the effect of reducing the growth rate of capital[7].

3.5   Variations and Extensions of the Model

In this section, I briefly introduce and discuss alternative specification of the model developed in sections 3.3 and 3.4. The purpose is to convince the reader that the results of the model are due to the interaction between asymmetric information incomplete contracting and that the results of the model are robust to modifications of the model.

---

7. In Appendix 3.5 I show the results for a version of the model where the growth rate of capital is exogenous and where the marginal value of capital is always increasing in ϕ for all experimented parameterizations.

(a) Beliefs

(b) Risky Assets to Total Capital
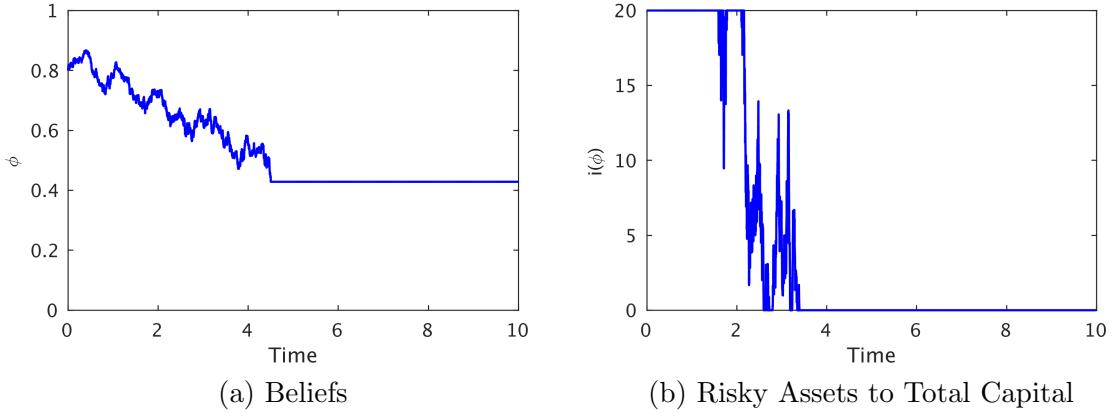
Figure 3.3: Bad Intermediary. Time series of investors's beliefs and demand for risky assets as share of total capital under management when returns are generated by the bad type. Parameter values are : $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

3.5.1   Model with Commitment

Consider a model that is identical to the model of section 3.3, but assume that the intermediary can commit to some portfolio risk i_t at each time t in a sequential game with investors. The assumption about the unobservability of the type is maintained but, at, each time t, the investor proposes and commit to a portfolio risk.

Investors are still learning the type of the intermediary by looking at the performance its portfolio. However, the signal manipulation incentive of the intermediary does not exist anymore. Indeed, as the intermediary commit to a particular trading strategy, investors will set their learning rule in order to update beliefs solely on the basis of the performance of the risky part of the intermediary's portfolio. Therefore the continuation value is still linear in capital, $\tilde{V}(K, \phi) = K\tilde{g}(\phi)$, but the HJB

equation takes the form

$$
\begin{aligned}
\rho\tilde{g}(\phi) = \max_{i \in [\hat{i}, \bar{i}]} \Big\{ & i(\phi\mu_1 + (1-\phi)\mu_2 - r) + \\
& + \tilde{g}'(\phi)\phi(1-\phi)^2 \left(\frac{\mu_1 - \mu_2}{\sigma}\right)^2 + \\
& + \tilde{g}(\phi)[r + i(1-\phi)(\mu_1 - \mu_2)] + \\
& + \frac{1}{2}\,\tilde{g}''(\phi)\phi^2(1-\phi)^2 \left(\frac{\mu_1 - \mu_2}{\sigma}\right)^2 + \\
& + \tilde{g}'(\phi)\phi(1-\phi)(\mu_1 - \mu_2)i \Big\}
\end{aligned}
\tag{3.13}
$$

Denote with $\tilde{i}(\phi)$ the equilibrium risk in the intermediary's portfolio.

A proposition analogous to Proposition 3.3 and lemmas analogous to 3.2 and 3.5 hold also for $\tilde{g}(\phi)$ if a bounded, continuously differentiable bounded function $\tilde{g}$ exists that satisfies (3.13) in $(\bar{\phi}, 1)$ with boundary conditions $\tilde{g}(\bar{\phi}) = 0$ and $\tilde{g}(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r}$.

The following result, when compared with Proposition 3.4, highlights the importance of the lack of commitment in generating hoarding of riskless assets

Proposition 3.5. *Suppose that there exists a continuously differentiable bounded function $\tilde{g}$ satisfying (3.13) in $(\bar{\phi}, 1)$ with boundary conditions $\tilde{g}(\bar{\phi}) = 0$ and $\tilde{g}(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r}$. If the intermediary can commit to a state contingent trading strategy $\tilde{i}(\phi)$, then $\tilde{i}(\phi) = \bar{i}$ for all $\phi \in (\bar{\phi}, 1]$.*

Numerical results are provided in Figure 3.4 and it is clear that, once the incentives to manipulate the signal are absent, the intermediary will make the socially optimal choice. Qualitatively, the results are robust to changes in the parameters.

(a) Marginal Value of Capital

(b) Risky Assets to Total Capital
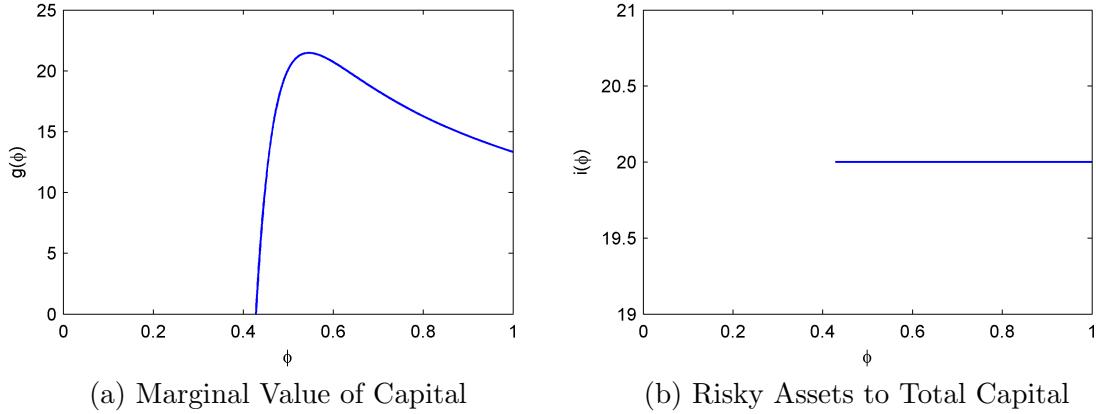
Figure 3.4: Commitment. Case 1. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

3.5.2 Exogenous Capital Growth

Consider now a situation where the intermediary is free from concerns about the growth rate of investors' capital. This could be the case of a small fund that is active in a large market with very diversified investors. Specifically, assume that the law of motion of the capital under management is

$$\frac{dK_t}{K_t} = \mu dt + \nu dZ_t \tag{3.14}$$

where Z_t is a Weiner process independent of W_t.

The fee will be determined by market clearing in the same way as before, so that expected return for the investors is equal to the risk free rate.

255

The HJB equation for the good type is now

$$\hat{\rho}\hat{g}(\phi) = \max_{0 \leq l \leq \bar{i}} \left\{ \hat{f}(\phi) + \right.$$
$$+ \hat{g}'(\phi)\frac{\varphi(\phi)}{\hat{i}(\phi)\sigma}[i((1-a)\mu_1 + a\mu_2 - r) - \hat{f}(\phi)] + \qquad (3.15)$$
$$\left. + \frac{1}{2}\hat{g}''(\phi)\frac{\varphi(\phi)^2}{\hat{i}(\phi)^2}i^2 \right\}$$

in the interval $(\bar{\phi}, 1]$ with terminal conditions $\hat{g}(\bar{\phi}) = 0$ and $\hat{g}(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r}$. In the previous equation, $\hat{\rho} = \rho - \mu$, $\hat{i}(\phi)$ is the equilibrium portfolio risk, defined in an analogous way to (3.11), and $\hat{f}(\phi) = \hat{i}(\phi)[\phi\mu_1 + (1-\phi)\mu_2 - r]$ is the equilibrium fee.

Results for a parameterization of the model are reported in Figure 3.5 and are robust to changes in the parameters. The qualitative features of $\hat{i}(\phi)$ are the same as in section 3.4. In fact, under existence conditions analogous to Assumption 3.4, Proposition 3.4 holds also for $\hat{g}(\phi)$ and $\hat{i}(\phi)$. However, from the numerical solutions, we can observe that the marginal value of capital, $\hat{g}(\phi)$, is always increasing in reputation. This provides suggestive evidence that the downward sloping part of the function $g(\phi)$ in section 3.3 is due to the decreasing rate of capital growth when $i(\phi) = \bar{i}$.

3.5.3 Heterogeneous Volatility

The results of the paper rely on the fact the two risky assets have the same volatility and that the two types choose the same portfolio risk. It is easy to see that the results continue to hold if the two assets have different volatility. Suppose that the

(a) Marginal Value of Capital (b) Risky Assets to Total Capital
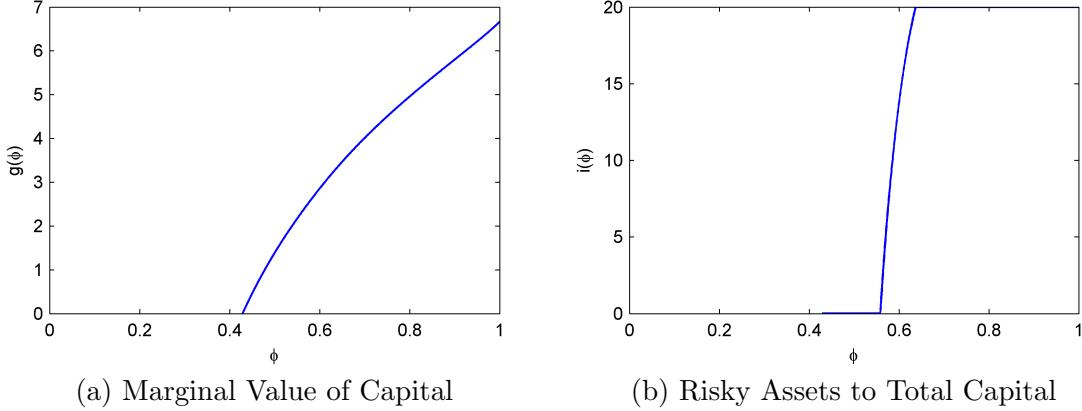
Figure 3.5: Exogenous Capital Growth. Case 1. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

good asset has volatility σ, while the bad asset has volatility $\frac{1}{k}\sigma$, for $k > 0$. This means that, in equilibrium $i_t^B = ki_t^G$ for every t, as a straightforward modification of Proposition 3.1. To simplify notation, let i_t be the equilibrium choice of the good type.

This implies that the expected return on the portfolio of the bad intermediary is $r + i_t k(\mu_2 - r)$, while the volatility is σi_t. But this is mathematically and economically equivalent to a model where both assets have the same volatility σ and the expected return of the bad asset is $\mu_2' = k(\mu_2 - r) + r$, which is still lower than r, since $\mu_2 - r < 0$. Therefore, heterogeneous volatility does not pose any threat to the conclusions of the model.

3.5.4 Time Varying Hidden Types

In a model like the one presented in sections 3.3 and 3.4, beliefs converge with probability one to either 1 or $\bar{\phi}$ depending on the type of the intermediary. This implies that episodes of liquidity hoarding will disappear in the long run. A natural extension is to consider a case in which the type of the intermediary, rather than being a time-constant process, evolves over time. The definitions of section 3.3 can be easily modified to explicitly deal with a process $(h_t)_{t \geq 0}$ for the type of the intermediary. However, Proposition 3.1 cannot be straightforwardly extended to this setting since, now, there is no guarantee that the continuation value of the bad intermediary is zero under perfect information. In this section, I provide a brief discussion of this model, which will be subject to further research in the future.

Suppose that $(h_t)_{t \geq 0}$ is independent of $(W_t)_{t \geq 0}$ and that it follows a continuous-time Markov chain with generator

$$\Lambda = \begin{pmatrix} -\lambda_2 & \lambda_2 \\ \lambda_1 & -\lambda_1 \end{pmatrix}$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$ and the state vector is $(1, 0)'$.

The law of motion of beliefs, conditional on $\theta_t > 0$ and on the two types pooling, is then modified as follows (see Theorem 9.1 in Liptser and Shiryaev, 2001):

$$d\phi_t = [-\lambda_2 \phi_t + \lambda_1 (1 - \phi_t)]dt + \varphi(i_t, \phi_t) \, (i_t \sigma)^{-1} \, (dR_t - \mu(i_t, \phi_t)dt) \, ;$$

while $d\phi_t = [-\lambda_2 \phi_t + \lambda_1 (1 - \phi_t)]dt$ when the intermediary does not generate signals.

At $\phi_t = 1$, reputation is deterministically decreased by an amount $-\lambda_2 dt$, while at $\bar{\phi}$ two cases have to be distinguished, depending on the stationary probability of the good type,

$$\pi = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

When $\pi \leq \bar{\phi}$, the reputation of the intermediary deterministically decreases when no signal is given to investors. Therefore, when the value $\bar{\phi}$ is hit, the intermediary will be shut down. On the contrary, when $\pi > \bar{\phi}$ the process ϕ_t deterministically moves upward when $\bar{\phi}$ is hit and, therefore, the continuation value of the intermediary at that point will not, in general, be zero. These two situations give rise to different terminal conditions, different continuation values when $\phi \in [0, \bar{\phi}]$ as well as possibly different incentives for the bad intermediary to imitate the good one or to separate itself.

In this richer framework, it is necessary to explicitly model the continuation value of both types in the interval $[\bar{\phi}, 1]$, conditional on them pooling. For the good type

we have

$$
\rho g(\phi) = \max_{0 \leq l \leq \bar{i}} \left\{ f(\phi) + \right.
$$

$$
+ g'(\phi) \frac{\varphi(\phi)}{i(\phi)\sigma} [i(\mu_1 - r) - f(\phi)] +
$$

$$
+ g(\phi)(r + i(\mu_1 - r) - f(\phi)) +
$$

$$
+ \frac{1}{2} g''(\phi) \frac{\varphi(\phi)^2}{i(\phi)^2} i^2 +
$$

$$
+ g'(\phi) \frac{\varphi(\phi)}{i(\phi)} i^2 \sigma +
$$

$$
+ \underbrace{g'(\phi)[-\lambda_2 \phi + \lambda_1(1 - \phi)]}_{\text{deterministic trend of } d\phi} +
$$

$$
\left. + \underbrace{\lambda_2[b(\phi) - g(\phi)]}_{\text{type's jump}} \right\}.
$$

(3.16)

and similarly, for the bad type,

$$
\rho b(\phi) = f(\phi) +
$$

$$
- b'(\phi)\phi^2(1 - \phi) \left(\frac{\mu_1 - \mu_2}{\sigma} \right)^2 +
$$

$$
+ b(\phi)(r - i(\phi)\phi(\mu_1 - \mu_2)) +
$$

$$
+ \frac{1}{2} b''(\phi)\phi^2(1 - \phi)^2 \left(\frac{\mu_1 - \mu_2}{\sigma} \right)^2 +
$$

$$
+ b'(\phi)\varphi(\phi)i(\phi)\sigma +
$$

$$
+ b'(\phi)[-\lambda_2 \phi + \lambda_1(1 - \phi)] +
$$

$$
+ \lambda_2[g(\phi) - b(\phi)].
$$

(3.17)

$i(\phi)$ is the equilibrium portfolio risk that such that, given $i(\phi)$ in equation (3.16),

the good intermediary optimally chooses $i = i(\phi)$ and the bad intermediary imitates it.

Consider first the case of $\pi \leq \bar{\phi}$, so that $\bar{\phi}$ is an absorbing barrier. Then, Proposition 3.1 holds and the bad intermediary imitates the good one. Indeed, if the two types were separated and the intermediary was bad, then beliefs would be reset to 0 and would deterministically converge to $\phi \leq \bar{\phi}$. Under the investors' perspective, expected returns are below the risk-free rate and therefore the intermediary will be inactive forever. This, therefore, provides us with boundary conditions $g(\bar{\phi}) = 0$ and $b(\bar{\phi}) = 0$. Moreover, for all $\phi \in [0, \bar{\phi}]$ we will have $g(\phi) = 0$ and $b(\phi) = 0$.

At $\phi = 1$, no learning happens by observing returns, but ϕ will deterministically decrease. Since the choice of i now affects only the growth rate of capital, in equilibrium we will have $i(1) = \bar{i}$ and the relations between $g(1)$, $g'(1)$, $b(1)$ and $b'(1)$ will be given by

$$g(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r} - g'(1)\frac{\lambda_2}{\rho - r} + [b(1) - g(1)]\frac{\lambda_2}{\rho - r} \tag{3.18}$$

and

$$b(1) = \frac{\bar{i}(\mu_1 - r)}{\rho - r + \bar{i}(\mu_1 - \mu_2)} - b'(1)\frac{\lambda_2}{\rho - r + \bar{i}(\mu_1 - \mu_2)} + [g(1) - b(1)]\frac{\lambda_1}{\rho - r + \bar{i}(\mu_1 - \mu_2)} \tag{3.19}$$

A numerical result is shown in Figure 3.6. Note that the parameterization differs from Figure 3.1 in section 3.4, but it can be compared with Figure 3.12 in Appendix 3.B. The qualitative features of the model are unaffected. However, the flight to riskless assets seems to be more dramatic, since the region of the state space where
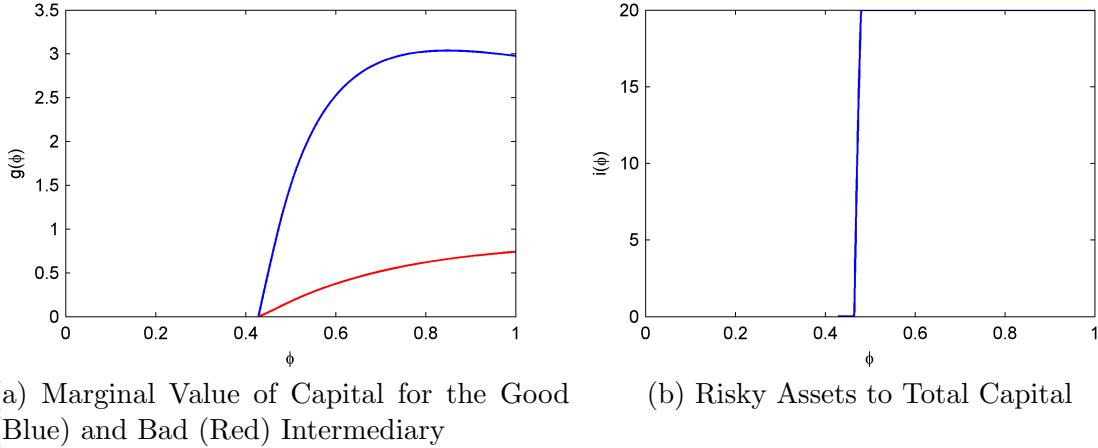
(a) Marginal Value of Capital for the Good (Blue) and Bad (Red) Intermediary

(b) Risky Assets to Total Capital

Figure 3.6: Switching Types. $\delta < \bar{\phi}$. Case 6. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.04$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

i shifts from $\bar{i}$ to $\hat{i}$ is now narrower.

Consider now the case of $\pi > \bar{\phi}$, so that, at $\bar{\phi}$, investors beliefs are deterministically move upward. The boundary conditions at $\bar{\phi}$, provided that the two types are pooling, are

$$g(\bar{\phi}) = g'(\bar{\phi})\frac{-\lambda_2\bar{\phi} + \lambda_1(1 - \bar{\phi})}{\rho - r} + [b(\bar{\phi}) - g(\bar{\phi})]\frac{\lambda_2}{\rho - r} \tag{3.20}$$

and

$$b(\bar{\phi}) = b'(\bar{\phi})\frac{-\lambda_2\bar{\phi} + \lambda_1(1 - \bar{\phi})}{\rho - r} + [g(\bar{\phi}) - b(\bar{\phi})]\frac{\lambda_1}{\rho - r}. \tag{3.21}$$

At $\phi = 1$, the boundary conditions are the same as in (3.18) and (3.19), provided that the two types are pooling.

262

The formal investigation of the properties of this model is still ongoing, but I can provide a result that holds under reasonable conditions. First, we need an analogous of Assumption 3.4, i.e. we need to assume the existence of bounded, twice continuously differentiable functions g and b that solve (3.16) and (3.17) – for an $i(\phi)$ such that the maximizer in (3.16) coincides with $i(\phi)$ – and that satisfy the boundary conditions (3.18), (3.19), (3.20) and (3.21). Call this condition C1. Second, I will need to show that $g(\bar{\phi}) \geq b(\bar{\phi})$ and $g(\phi) \geq b(\phi)$ for ϕ is a right-neighborhood of $\bar{\phi}$. Let this be condition C2.

Claim 3.1. *If conditions C1 and C2 hold, then there exists an $\epsilon > 0$ such that $g''(\phi) < 0$ and $i(\phi) = \hat{i}$ for all $\phi \in (\bar{\phi}, \bar{\phi} + \epsilon)$.*

The proof is omitted, since it follows the same reasoning of the proof of Proposition 3.4, after noting two facts. The first one is that $g'(\bar{\phi}) > 0$ follows directly from C2 and (3.20). The second one is that, at least in a right-neighborhood of $\bar{\phi}$, the bad intermediary will pool, for, otherwise $g(\bar{\phi}) = g(0)$ and, since the intermediary receives no flow payoff when $\phi \in [0, \bar{\phi}]$, this would imply that $g(\bar{\phi}) = 0$.

The fact that Claim 3.1 holds when $\pi > \bar{\phi}$ underscores the intuition supporting that the value of the signaling option is a key driver of the model's results. Indeed, if reputation falls below the threshold, the good intermediary would be unable to signal its type until $\bar{\phi}$ is reached again. This option may be less valuable if $\pi < \bar{\phi}$. This is because the deterministic drift of beliefs may be negative enough that, even under the information filtration of the good type, beliefs may still be drifting downward.

3.6   Conclusions

I developed a model of financial intermediation with asymmetric information that delivers outcomes that qualitatively resemble episodes of liquidity hoarding. The incentives of the intermediary to manipulate the evolution of investors' beliefs, coupled with the impossibility to commit to a trading strategy, are the key drivers of the results. When reputation is close to the threshold below which the intermediary is terminated, a risk neutral intermediary displays risk aversion, despite being subject to a convex compensation scheme. This is because, in an infinite horizon game, termination would make the good intermediary lose the option to signal its type and gain reputation.

Appendix 3.A   Proofs

3.A.1   Proof of Lemma 3.1

Here, I provide a lemma supporting the claim that restricting our attention to public strategies is without loss of generality.

Let $(\mathcal{F}_t^P)_{t \geq 0}$ be the filtration generated by h and $(R_t)_{t \geq 0}$ and is possibly augmented by the collection of P-null sets. Recall that we are dealing with a probability space $(\Omega, \mathcal{F}^*, P)$ and that I have defined $(\mathcal{F}_t)_{t \geq 0}$ as the filtration generated by $(W_t)_{t \geq 0}$ and h.

Lemma 3.1. *Given the assumptions that $\mathcal{F}_0 = \mathcal{F}_0^P$ and that the filtrations $(\mathcal{F}_t)_{t \geq 0}$ and $(\mathcal{F}_t^P)_{t \geq 0}$ are augmented by the P-null sets, we have that $\mathcal{F}_t = \mathcal{F}_t^P$ for all $t \geq 0$. Therefore, there is no loss of generality in assuming that strategies are public.*

Proof. Let us proceed by contradiction and suppose that there exists a τ defined as

$$\tau = \inf\{t \geq 0 : \exists B \text{ s.t. } B \in \mathcal{F}_t \text{ and } B \notin \mathcal{F}_t^P\}$$

and note that, for every t, $\mathcal{F}_t^P \subseteq \mathcal{F}_t$.

Since both filtrations are generated by continuous processes, then they are left-continuous (see Karatzas and Shreve (1991), Chapter 2.7). A filtration $(\mathcal{G}_t)_{t \geq 0}$ is said to be left-continuous if $\mathcal{G}_t = \mathcal{G}_{t-}$, where $\mathcal{G}_{t-} \equiv \sigma\left(\cup_{s<t}\mathcal{G}_s\right)$. By convention, $\mathcal{G}_{0-} = \mathcal{G}_0$.

Suppose then that $\mathcal{F}_\tau^P \subset \mathcal{F}_\tau$. By definition of τ we must have, $\mathcal{F}_{\tau-}^P = \mathcal{F}_{\tau-}$, leading us to a contradiction, since, by left-continuity, $\mathcal{F}_{\tau-}^P = \mathcal{F}_{\tau-} = \mathcal{F}_\tau$.

Suppose instead that $\mathcal{F}_\tau^P = \mathcal{F}_\tau$. By Proposition 7.7 in Chapter 2.7 of Karatzas and Shreve (1991), the filtration $(\mathcal{F}_t)_{t\geq 0}$ is also right-continuous. A filtration $(\mathcal{G}_t)_{t\geq 0}$ is said to be right-continuous if $\mathcal{G}_t = \mathcal{G}_{t+}$, where $\mathcal{G}_{t+} \equiv \cap_{s>t}\mathcal{G}_s$.

This means that we must have $\mathcal{F}_\tau = \mathcal{F}_{\tau+}$. However, by definition of τ we must also have $\mathcal{F}_{\tau+}^P \subset \mathcal{F}_{t+}$. This would imply that $\mathcal{F}_{\tau+}^P \subset \mathcal{F}_\tau^P$, contradicting that $(\mathcal{F}_t^P)_{t\geq 0}$ is an increasing sequence of σ-algebras.

We therefore conclude that $\mathcal{F}_t^P = \mathcal{F}_t$ for all t. $\qquad\qquad\square$

3.A.2   Proof of Proposition 3.1

This proof follows the same lines of Theorem 7.17 and Lemma 5.2 in Liptser and Shiryaev (2001) with some minor modifications to allow the proof to work for any finite time t in an infinite horizon game.

Proof. Given a time interval $[0, t]$, consider a partition of n sub-intervals with end points $0 = t_{n,0} < t_{n,1} < \cdots < t_{n,n} = t$ such that $\max_i |t_{n,i+1} - t_{n,i}| \to 0$ as $n \to \infty$.

Then consider

$$
\sum_{i=0}^{n-1} (R_{t_{n,i+1}} - R_{t_{n,i}})^2 =
$$

$$
\sum_{i=0}^{n-1} \left(\int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma_s dW_s \right)^2 +
$$

$$
+ \sum_{i=0}^{n-1} \left\{ \int_{t_{n,i}}^{t_{n,i+1}} [r + i_s(h\mu_1 + (1-h)\mu_2 - r)]ds \right\}^2 \qquad (3.22)
$$

$$
+ \sum_{i=0}^{n-1} 2 \left(\int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma_s dW_s \right) \left\{ \int_{t_{n,i}}^{t_{n,i+1}} [r + i_s(h\mu_1 + (1-h)\mu_2 - r)]ds \right\}
$$

Note that $|r + i_s(h\mu_1 + (1-h)\mu_2 - r)| \le r + \bar{i}D$, for $D \equiv |\mu_1 - r| + |\mu_2 - r|$,

therefore

$$\sum_{i=0}^{n-1} \left\{ \int_{t_{n,i}}^{t_{n,i+1}} [r + i_s(h\mu_1 + (1-h)\mu_2 - r)]ds \right\}^2$$

$$\le (r + \bar{i}D)^2 \cdot t \cdot \max_i |t_{n,i+1} - t_{n,i}| \to 0, \qquad \text{as } n \to \infty.$$

Also,

$$\left| \sum_{i=0}^{n-1} \left(\int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma_s dW_s \right) \left\{ \int_{t_{n,i}}^{t_{n,i+1}} [r + i_s(h\mu_1 + (1-h)\mu_2 - r)]ds \right\} \right| \le$$

$$\le (r + \bar{i}D) \cdot t \cdot \max_i \left| \int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma_s dW_s \right| \to 0 \quad \text{as } n \to \infty.$$

since $|t_{n,i+1} - t_{n,i}| \to 0$ for all i and $\{i_s\sigma\}_{u\ge0}$ is a bounded process.

As for the remaining term, note that, by Ito's lemma

$$d\left(\int_{t_{n,i}}^{s} i_u \sigma \, dW_u \right)^2 = 2\left(\int_{t_{n,i}}^{s} i_u \sigma \, dW_u \right) i_s \sigma dW_s + (i_s\sigma)^2 ds$$

and hence

$$\left(\int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma \, dW_s \right)^2 = 2 \int_{t_{n,i}}^{t_{n,i+1}} \left(\int_{t_{n,i}}^{s_{n,i}} i_u \sigma \, dW_u \right) i_{s_{n,i}} \sigma dW_{s_{n,i}} + \int_{t_{n,i}}^{t_{n,i+1}} (i_s\sigma)^2 ds,$$

with $s_{n,1} \in [t_{n,i}, t_{n+1,i}]$.

We can then substitute the latter expression in the first term in the right-hand

side of (3.22) and obtain

$$\sum_{i=0}^{n-1} \left(\int_{t_{n,i}}^{t_{n,i+1}} i_s \sigma_s dW_s \right)^2 = \int_0^t (i_s \sigma)^2 ds + 2 \int_0^t \left(\int_{t_{n,i}}^s i_u \sigma \, dW_u \right) i_s \sigma dW_s. \quad (3.23)$$

The second term in the right-hand side of (3.23), however, converges in probability to zero, since

$$\left| \int_{t_{n,i}}^{t_{n,i+1}} \left(\int_{t_{n,i}}^{s_{n,i}} i_u \sigma \, dW_u \right) i_s \sigma dW_s \right|$$

$$\leq \left[\max_i \sup_{s_{n,1} \in [t_{n,i}, t_{n+1,i}]} \left(\int_{t_{n,i}}^{s_{n,i}} i_u \sigma \, dW_u \right)^2 \right] \cdot (\bar{i} \sigma)^2 \cdot t \to 0$$

$$\text{as } n \to \infty$$

since $|t_{n,i+1} - t_{n,i}| \to 0$ for all i and $\{i_u \sigma\}_{u \geq 0}$ is a bounded process.

The left-hand side of (3.22) is clearly $\mathcal{F}_t^I$-measurable and converges to an $\mathcal{F}_t^I$-measurable random variable as $n \to \infty$, while the right-hand side converges in probability to $\int_0^t (i_s \sigma)^2 \, ds$. It therefore follows that $\int_0^t (i_s \sigma)^2 \, ds$ is itself $\mathcal{F}_t^I$-measurable.

We can then straightforwardly apply Lemma 5.2 in Liptser and Shiryaev (2001) and use the fact that both i_t and σ are strictly positive, to conclude that there exist an $\mathcal{F}_t^I$-measurable process $(\tilde{i}_s)_{0 \leq s \leq t}$, that is progressively measurable with respect to $(\mathcal{F}_s^I)_{0 \leq s \leq t}$ and such that $\tilde{i}_t = i_t$ P-a.s..

Therefore, let $(i_t^G)_{t \geq 0}$ be the equilibrium public strategy of the good type, then

consistency of beliefs imposes

$$\phi_t = 0 \quad \text{if} \quad \exists s \geq 0, \Delta > 0 \text{ s.t.} \quad \int_s^{s+\Delta} (\tilde{i}_u \sigma)^2 \, du \neq \int_s^{s+\Delta} (i_u^G \sigma)^2 \, du.$$

Therefore, it is optimal for the bad intermediary to pool P-a.s.

$\square$

3.A.3   Proof of Proposition 3.2

In equilibrium, the law of motion of the signal is given by

$$dR_t = [r + i_t(h\mu_1 + (1-h)\mu_2 - r)]dt + i_t \sigma dW_t$$

where $(i_t)_{t\geq 0}$ is an $(\mathcal{F}_t^I)_{t\geq 0}$ progressively measurable process, while h is hidden to investors. The fact that only the drift changes depending on the type of the intermediary calls for an application of Girsanov's theorem to find an appropriate change of measure and derive the likelihood ratio of the two types. However, care should be taken since Girsanov's theorem cannot be extended straightforwardly to the infinite horizon case. Fortunately, Huang and Pages (1992) and Revuz and Yor (2013) provide a framework to extend the result to the infinite horizon case.

Proof. Let P^G be a probability measure on $(\Omega, \mathcal{F}_\infty^I)$ induced on the good type. Define P^B in an analogous way for the bad type.

We are interested in finding a random variable ξ_t representing the ratio between the likelihood that the path $(R_s)_{0\leq s\leq t}$ is generated by the good type and

the likelihood that the same path is generated by the bad type. Note that $\tilde{W}_t \equiv \int_0^t [dR_s - rds - i_s(\mu_2 - r)](i_s \sigma)^{-1} ds$ is a Weiner process conditional on the bad type.

Let

$$\eta = (\mu_1 - \mu_2)\sigma^{-1}$$

and define the local martingale

$$\xi_t = \exp\left\{ \eta \tilde{W}_t - \frac{1}{2}\eta^2 t \right\}.$$

It can be seen that $\mathrm{E}_{PB}[\xi_t] = 1$ for all t. From Proposition 1 in Huang and Pages (1992) it follows that there exists a measure Q on $(\Omega, \mathcal{F}_\infty^I)$ such that the restriction of Q on $\mathcal{F}_t^I$ is equivalent to the restriction of P^B on $\mathcal{F}_t$ and, moreover, if restricted to $(\Omega, \mathcal{F}_t^I)$, ξ_t is the likelihood ratio dQ/dP^B. Furthermore,

$$\tilde{W}_t^Q = \tilde{W}_t - \eta t$$

is a standard Weiner process for $(\Omega, \mathcal{F}_\infty^I, Q)$. We can think of Q as a probability measure whose restriction on $(\Omega, \mathcal{F}_t^I)$ coincides with the restriction on $(\Omega, \mathcal{F}_t^I)$ of P^G, for all t.

By consistency of beliefs in equilibrium

$$\phi_t = \frac{p\xi_t}{p\xi_t + (1-p)} \tag{3.24}$$

If we then apply Ito's lemma we obtain

$$d\phi_t = -\frac{(1-p)p^2}{(p\xi_t + (1-p))^3}(\eta\xi_t)^2 dt + \frac{(1-p)p}{(p\xi_t + (1-p))^2}\eta\xi_t d\tilde{W}_t.$$

Finally recall that $d\tilde{W}_t = [dR_s - rds - i_s(\mu_2 - r)ds](i_s\sigma)^{-1}$. It is then sufficient to substitute for $d\tilde{W}_t$ in the previous expression and use (3.24) to conclude that

$$d\phi_t = (1-\phi_t)\phi_t\frac{\mu_1 - \mu_2}{\sigma}(i_t\sigma)^{-1}(dR_t - \mu(i_t,\phi_t)dt)$$

where $\mu(i_t,\phi_t) = r + i_t\left[\phi_t\mu_1 + (1-\phi_t)\mu_2 - r\right]$. $\qquad\square$

3.A.4   Proof of Proposition 3.3

The proof of this Proposition follows the same lines of standard verification theorems as those in Chapter 10 of Oksendal (2013). However, given the game-theoretic set-up of the model, it is essential to keep in mind that, given an equilibrium portfolio risk $i(\phi)$, the good intermediary is free to choose its portfolio risk. The two will coincide in equilibrium by the consistency condition given by (3.11). Expectations are understood to be expectations conditional on $\mathcal{F}_0$.

Proof. To begin with, note that $V(0,\phi) = 0$ and the function $V(K,\phi) = Kg(\phi)$ satisfies this condition. Similarly, $V(K,\phi) = 0$ if $\phi \leq \bar{\phi}$ and $Kg(\phi)$ satisfies also this condition. Let us therefore focus on the case $K_0 > 0$ and $\phi_0 > \bar{\phi}$.

Define $\tau = \inf\{t : \phi_t \le \bar\phi\}$. I want to show that

$$K_0 g(\phi_0) \ge \mathrm{E}\left[\int_0^\tau e^{-\rho t} f(\phi_t) K_t dt\right]$$

for any feasible process of the control $(i_t)_{t\ge 0}$ when K_t follows (3.1), ϕ_t follows (3.4) and $f(\phi)$ and $i(\phi)$ are defined by (3.6) and (3.11).

Let

$$W_t = e^{-\rho \min\{t,\tau\}} K_{\min\{t,\tau\}} g(\phi_{\min\{t,\tau\}}) + \int_0^{\min\{t,\tau\}} e^{-\rho s} f(\phi_s) K_s \, ds \quad .$$

Define

$$\mathcal{D}[Kg(\phi), i] = \left\{ f(\phi) + g'(\phi)\varphi(\phi)\frac{[i(\mu_1 - r) - f(\phi)]}{i(\phi)\sigma} + g(\phi)(-\rho + r + i(\mu_1 - r) - f(\phi)) + \right.$$
$$\left. + \frac{1}{2}g''(\phi)\varphi(\phi)^2 \frac{i^2}{i(\phi)^2} + g'(\phi)\varphi(\phi)\sigma\frac{i^2}{i(\phi)} \right\} K$$

and

$$\mathcal{V}[Kg(\phi), i] = \left\{ g'(\phi)\varphi(\phi)\frac{i}{i(\phi)} + g(\phi)\sigma i \right\} K.$$

By Ito's lemma

$$W_t = W_0 + \int_0^{\min\{t,\tau\}} e^{-\rho s}\mathcal{D}[K_s g(\phi_s), i_s]\,ds + \int_0^{\min\{t,\tau\}} e^{-\rho s}\mathcal{V}[K_s g(\phi_s), i_s]\,W_s.$$

$$(3.25)$$

By (3.10) and (3.11), $\mathcal{D}[K_s g(\phi_s), i_s]$ attains its maximum when $i_s = i(\phi_s)$ and

272

$\mathcal{D}[K_s g(\phi_s), i_s] \leq 0$ for any i_s. Therefore

$$\int_0^{\min\{t,\tau\}} e^{-\rho s} \mathcal{V}[K_s g(\phi_s), i_s] W_s \geq W_t - W_0. \tag{3.26}$$

with equality if the control $(i_s)_{s \geq 0}$ coincides with the optimal one.

The left-hand side of (3.26) is a local martingale that is bounded from below by $-W_0 = -V_0 > -\infty$. Hence, it is a supermartingale. Taking expectations in (3.25) we obtain

$$\mathrm{E}[W_t] \leq W_0 + \mathrm{E}\left[\int_0^{\min\{t,\tau\}} e^{-\rho s} \mathcal{D}[K_s g(\phi_s), i_s]\, ds\right] \leq W_0 = K_0 g(\phi_0)$$

with equalities if the control $(i_s)_{s \geq 0}$ is the optimal one.

Hence,

$$K_0 g(\phi_0) \geq \mathrm{E}\left[\int_0^{\min\{t,\tau\}} e^{-\rho s} f(\phi_s) K_s\, ds\right] + \mathrm{E}\left[e^{-\rho \min\{t,\tau\}} K_{\min\{t,\tau\}} g(\phi_{\min\{t,\tau\}})\right].$$

Using the fact that we must have

$$\lim_{t \to \infty} \mathrm{E}\left[e^{-\rho \min\{t,\tau\}} K_{\min\{t,\tau\}} g(\phi_{\min\{t,\tau\}})\right] = 0.$$

then, by the monotone convergence theorem, as $t \to \infty$, we conclude

$$K_0 g(\phi_0) \geq \mathrm{E}\left[\int_0^\tau e^{-\rho s} f(\phi_s) K_s\, ds\right]$$

for any feasible portfolio risk process $(i_s)_{s \geq 0}$, with equality if $i_s = i(\phi_s)$ P-a.s..

273

It follows that $K_0 g(\phi_0)$ is the continuation value and $(i(\phi_t))_{t \geq 0}$ is indeed the strategy played by the intermediary in equilibrium.

$\square$

3.A.5   Proof of Proposition 3.4

Let us re-write (3.12) with a more parsimonious notation as

$$\frac{1}{2} g''(\phi)\phi^2(1-\phi)^2\kappa^2 =$$

$$-i(\phi)(-\delta_2+\phi\delta_1) - g'(\phi)\phi(1-\phi)^2\kappa^2 + g(\phi)(\delta - i(\phi)(1-\phi)\delta_1) - g'(\phi)\phi(1-\phi)\delta_1 i(\phi)$$

$$(3.27)$$

where the parameters $\kappa, \delta_1, \delta_2$ and δ are all strictly positive.

I break down the proof in a series of Lemmas. Assumption 3.4 is taken for granted in all of them. All limits in the form $\phi \to \bar{\phi}$ are understood to be for sequences converging to $\bar{\phi}$ from above.

Lemma 3.2. $\liminf_{\phi \to \bar{\phi}} g'(\phi) \geq 0$.

Proof. Suppose, towards a contradiction, that $\liminf_{\phi \to \bar{\phi}} g'(\phi) < 0$. Then, in any interval $(\bar{\phi}, \bar{\phi} + \varepsilon)$, there exists a sequence $\phi_n \to \bar{\phi}$ such that $\lim_{n\to\infty} g'(\phi_n) < 0$. Let $\varepsilon_n = \phi_n - \bar{\phi}$. By a first order Taylor expansion

$$g(\bar{\phi}) = g(\phi_n) - g'(\phi_n)\varepsilon_n + o(\varepsilon_n) \implies g(\phi_n) = \left(g'(\phi_n) + \frac{o(\varepsilon_n)}{\varepsilon_n}\right)\varepsilon_n$$

Then, there is an $\bar{n} > 0$ large enough such that $\left(g'(\phi_{\bar{n}}) + \frac{o(\varepsilon_{\bar{n}})}{\varepsilon_{\bar{n}}}\right)\varepsilon_{\bar{n}} < 0$, which

contradicts the fact that $g(\phi_{\bar{n}})$ must be non-negative.

$\square$

Lemma 3.3. *There exists an $\varepsilon > 0$ such that, for all $\phi \in (\bar{\phi}, \bar{\phi} + \varepsilon)$, $g''(\phi) \leq 0$.*

Proof. By lemma 3.2 and by equation (3.27) we must conclude that $\limsup_{\phi \to \bar{\phi}} g''(\phi) \leq 0$, that is, there exists an $\varepsilon > 0$ such that, for all $\phi \in (\bar{\phi}, \bar{\phi} + \varepsilon)$ we have that $g''(\phi) \leq 0$.

$\square$

Lemma 3.4. $\liminf_{\phi \to \bar{\phi}} g'(\phi) = \limsup_{\phi \to \bar{\phi}} g'(\phi)$ *and possibly both are equal to* $+\infty$. *Moreover,* $\liminf_{\phi \to \bar{\phi}} g''(\phi) = \limsup_{\phi \to \bar{\phi}} g''(\phi)$ *and possibly both are equal to* $-\infty$.

Proof. Recall that, by Lemmas 3.2 and 3.3, there exists a neighborhood of $\bar{\phi}$ where $\liminf_{\phi \to \bar{\phi}} g'(\phi) \geq 0$ and $\limsup_{\phi \to \bar{\phi}} g''(\phi) \leq 0$ for all ϕ in that neighborhood.

Suppose, by way of contradiction, that $\liminf_{\phi \to \bar{\phi}} g'(\phi) = H \geq 0$ and that $\limsup_{\phi \to \bar{\phi}} g'(\phi) > H$. This means that there exists an $H' > H$ such that $g'(\phi) = H'$ for infinitely many ϕ in any neighborhood of $\bar{\phi}$, contradicting the fact that $g''(\phi) \leq 0$ for all ϕ in a neighborhood of $\bar{\phi}$.

This, together with equation (3.27), implies the claim.

$\square$

With some abuse of notation, let $g'(\bar{\phi}) \equiv \lim_{\phi \to \bar{\phi}} g'(\phi)$ and $g''(\bar{\phi}) \equiv \lim_{\phi \to \bar{\phi}} g''(\phi)$. It is understood that both can be infinity.

Lemma 3.5. $g'(\bar{\phi}) > 0$ *and* $g''(\bar{\phi}) < 0$.

Proof. Suppose, by way of contradiction, that $g'(\bar{\phi}) = 0$. By (3.27), this also implies $g''(\bar{\phi}) = 0$. Let $\hat{\phi} \equiv \sup\{\phi : \phi \geq \bar{\phi}, g''(\phi) = 0\}$. We must have $g''(\phi) > 0$ for all $\phi \in (\hat{\phi}, \hat{\phi} + \varepsilon)$ for some small enough ε, for otherwise we would have $g(\phi) < 0$, since

$$g(\phi) = \int_{\hat{\phi}}^{\phi} \int_{\hat{\phi}}^{\phi} g''(x) dx dx$$

Consider $\varepsilon' \in (0, \varepsilon)$ and let $\phi' = \hat{\phi} + \varepsilon'$. By a first order Taylor expansion

$$g(\phi') = g(\hat{\phi}) + g'(\hat{\phi})\varepsilon' + o(\varepsilon) = o(\varepsilon) \tag{3.28}$$

$$g'(\phi') = g'(\hat{\phi}) + g''(\hat{\phi})\varepsilon' + o(\varepsilon) = o(\varepsilon) \tag{3.29}$$

Substituting (3.28) and (3.29) into (3.27) we obtain

$$\frac{1}{2}g''(\phi')(\phi')^2(1 - \phi')^2 \kappa^2 = \left(-\delta_1 i(\phi') + \frac{o(\epsilon')}{\epsilon'}\right)\varepsilon'$$

Since $\delta_1 > 0$ and $i(\phi')$ is strictly positive and bounded away from zero, it follows that, for ε' small enough, $g''(\phi') < 0$, contradicting that $g''(\phi) > 0$ for all $\hat{\phi} < \phi < \hat{\phi} + \varepsilon$.

Therefore it follows that $g'(\bar{\phi}) > 0$ and $g''(\bar{\phi}) < 0$.

$\square$

Lemma 3.6. $i(\phi) = \hat{i}$ *in a neighborhood of* $\bar{\phi}$.

Proof. An immediate corollary of the Lemmas of this part of the appendix is that

$$\frac{1}{2}g''(\phi)\frac{\varphi(\phi)}{i(\phi)^2} + g'(\phi)\frac{\varphi(\phi)\sigma}{i(\phi)} \leq 0$$

in a neighborhood of $\bar{\phi}$.

It follows that, in this neighborhood, the unconstrained maximizer in (3.10) is positive and equal to

$$M(\phi) = -\frac{g'(\phi)\varphi(\phi)\frac{\mu_1-r}{\sigma} + g(\phi)i(\phi)(\mu_1-r)}{g''(\phi)\frac{\varphi(\phi)^2}{i(\phi)} + 2g'(\phi)\varphi(\phi)\sigma}.$$

I want to show that $M(\phi) < i(\phi)$ if ϕ is close enough to $\bar{\phi}$ and, hence, that the only possible equilibrium is the one where $i(\phi) = \hat{i}$.

Towards a contradiction, suppose that, in equilibrium, $M(\phi) \geq i(\phi)$. It then follows that

$$\frac{1}{2}g''(\phi)\varphi(\phi) \geq -g'(\phi)\varphi(\phi)\sigma i(\phi) - \frac{1}{2}g'(\phi)\varphi(\phi)\frac{\mu_1-r}{\sigma} - \frac{1}{2}g(\phi)i(\phi)(\mu_1-r)$$

Substituting the latter expression in (3.12) we obtain

$$\rho g(\phi) \geq f(\phi) + g'(\phi)\frac{\varphi(\phi)}{\sigma}\left[\frac{1}{2}(\mu_1-r) - \frac{f(\phi)}{i(\phi)}\right] + g(\phi)\left[r + \frac{1}{2}i(\phi)(\mu_1-r) - f(\phi)\right].$$

Because $f(\phi) \to 0$ and $g(\phi) \to 0$ as $\phi \to \bar{\phi}$, we would conclude that $\lim_{\phi \to \bar{\phi}} g'(\phi) \leq 0$. But this contradicts Lemma 3.5. Therefore, for ϕ close enough to $\bar{\phi}$, we must have $M(\phi) < i(\phi)$. $\square$

3.A.6   Proof of Proposition 3.5

The right-hand side of equation (3.13) is linear in i. To prove the claim, it is therefore sufficient to show that, for all $\phi \in (\bar{\phi}, 1]$,

$$C(\phi) \equiv (\phi \mu_1 + (1 - \phi)\mu_2 - r) + \tilde{g}(\phi)(1 - \phi)(\mu_1 - \mu_2) + \tilde{g}'(\phi)\phi(1 - \phi)(\mu_1 - \mu_2) > 0.$$

Proof. Recall that an analogous of Proposition 3.4 holds for $\tilde{g}(\phi)$, so that, $\tilde{g}'(\bar{\phi}) > 0$. Since $\tilde{g}'(\bar{\phi}) > 0$, if ϕ is sufficiently close to $\bar{\phi}$, then $C(\phi) > 0$. By assumption, $C(\phi)$ is continuously differentiable in $(\bar{\phi}, 1)$. Therefore, consider $\phi' \equiv \inf\{\phi : \phi > \bar{\phi} \text{ and } C(\phi) \leq 0\}$. By way of contradiction, suppose such ϕ' exists. Then we must have $C(\phi') = 0$ and $C'(\phi') \leq 0$.

From $C(\phi') = 0$, I can derive that

$$-\tilde{g}(\phi') - \tilde{g}'(\phi')\phi' = \frac{(\phi'\mu_1 + (1 - \phi')\mu_2 - r)}{\phi'(1 - \phi')(\mu_1 - \mu_2)} > 0 \tag{3.30}$$

while from (3.13), we obtain

$$2\tilde{g}'(\phi')(1 - \phi') = 2\frac{(\rho - r)\tilde{g}(\phi')}{\phi'(1 - \phi')\kappa^2} - \tilde{g}''(\phi')\phi'(1 - \phi') \tag{3.31}$$

where $\kappa = \frac{\mu_1 - \mu_2}{\sigma}$.

Compute

$$\frac{C'(\phi')}{\mu_1 - \mu_2} = 1 + \tilde{g}'(\phi')(1 - \phi') - \tilde{g}(\phi') + \tilde{g}''(\phi')\phi'(1 - \phi') + \tilde{g}'(\phi')(1 - \phi') - \tilde{g}'(\phi')\phi'$$

and use (3.30) and (3.31) to obtain

$$\frac{C'(\phi')}{\mu_1 - \mu_2} = 1 - \tilde{g}(\phi') - \tilde{g}'(\phi')\phi' + 2\frac{(\rho - r)\tilde{g}(\phi')}{\phi'(1 - \phi')\kappa^2} > 0$$

which contradicts that $C'(\phi') \leq 0$.

$\square$

Appendix 3.B   Additional Plots

(a) Euro Area   (b) United States
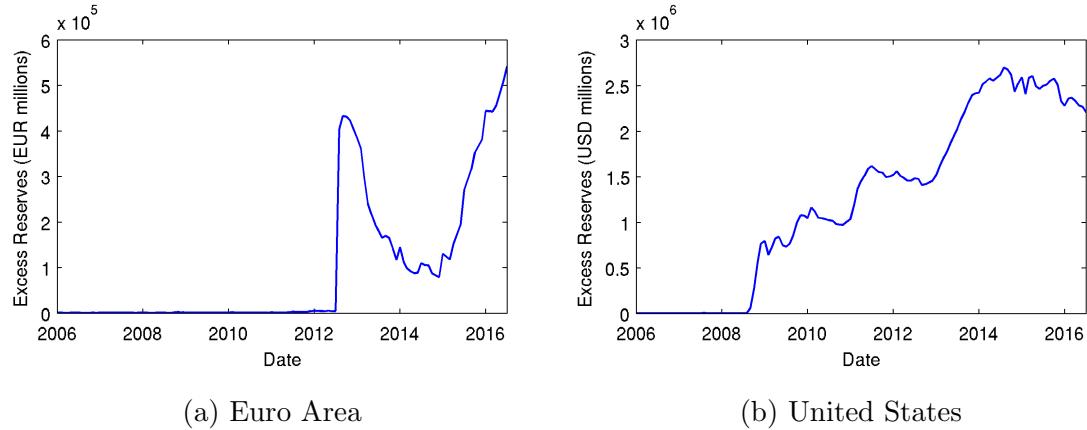
Figure 3.7: Excess reserves of institutions subject to minimum reserve requirements in the Euro Area and the United States
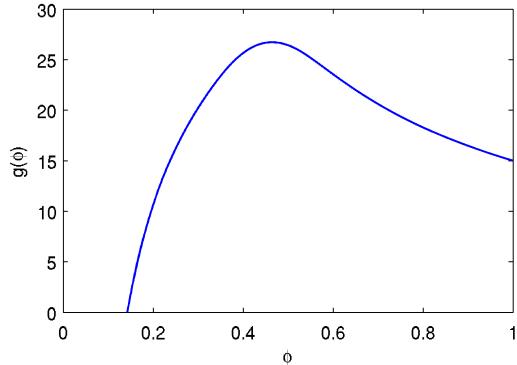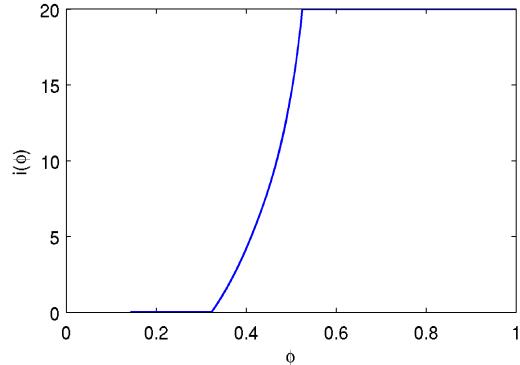
(a) Marginal Value of Capital　　　　(b) Risky Assets to Total Capital

Figure 3.8: Case 2. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.02$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.
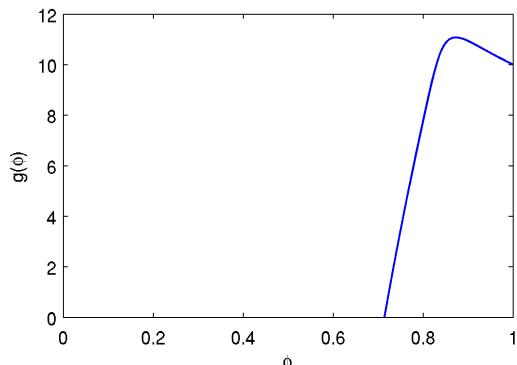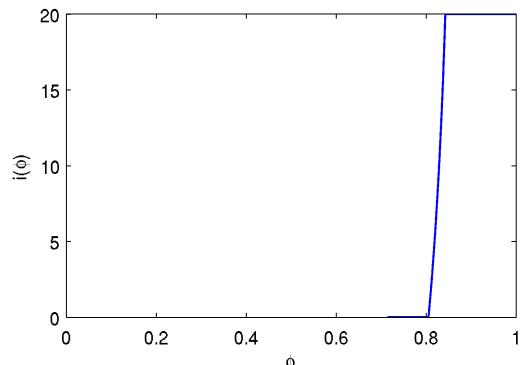

(a) Marginal Value of Capital　　　　(b) Risky Assets to Total Capital

Figure 3.9: Case 3. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.04$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.
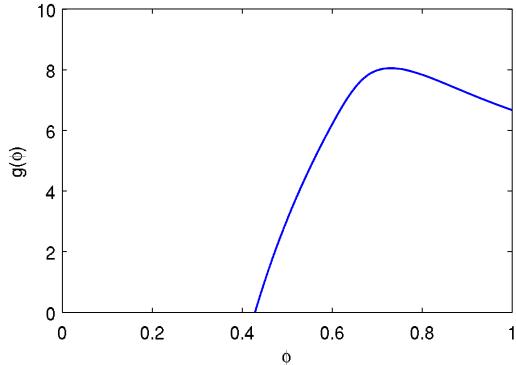
(a) Marginal Value of Capital
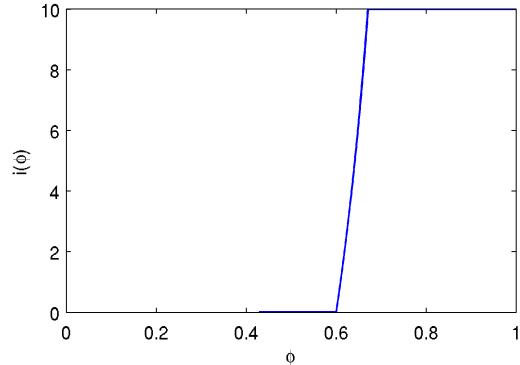
(b) Risky Assets to Total Capital

Figure 3.10: Case 4. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 10$ and $\hat{i} = 10^{-9}$.
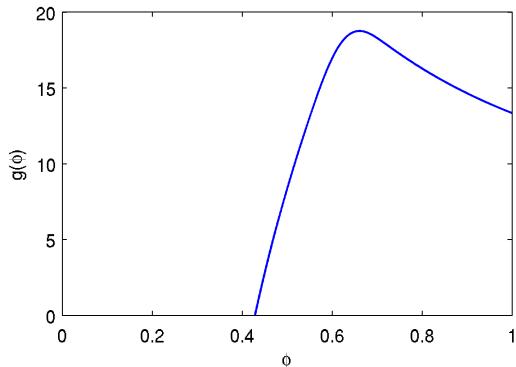

(a) Marginal Value of Capital
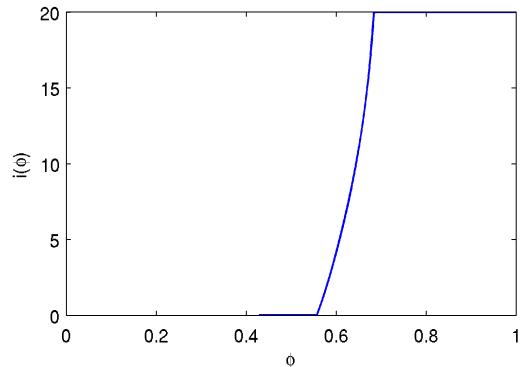
(b) Risky Assets to Total Capital

Figure 3.11: Case 5. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.10$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

281

(a) Marginal Value of Capital

(b) Risky Assets to Total Capital
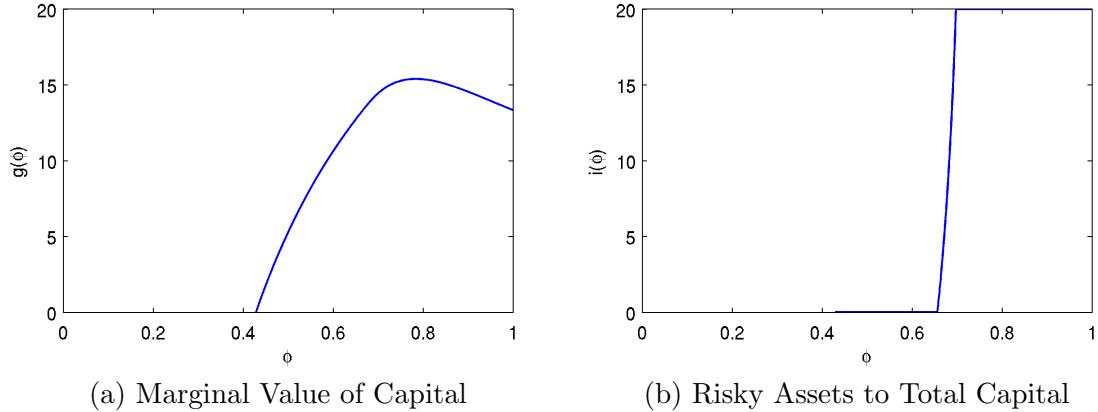
Figure 3.12: Case 6. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.04$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.


(a) Marginal Value of Capital

(b) Risky Assets to Total Capital
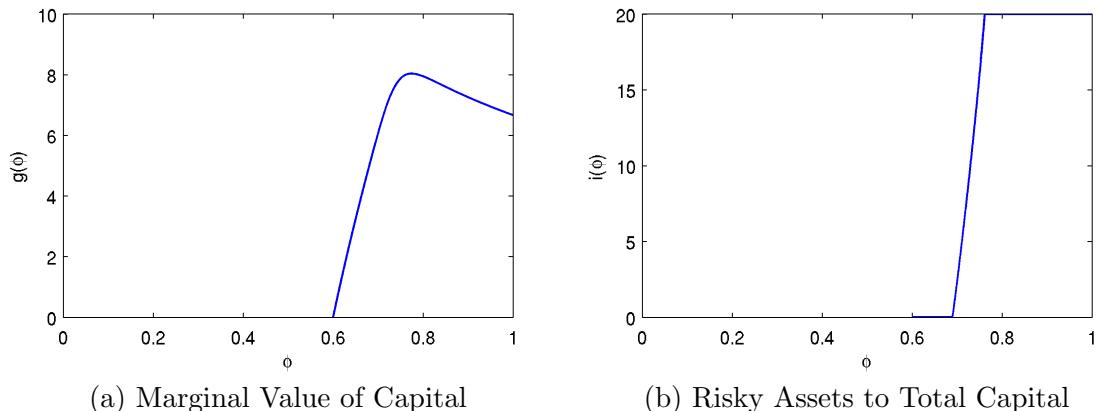
Figure 3.13: Case 7. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.04$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

282
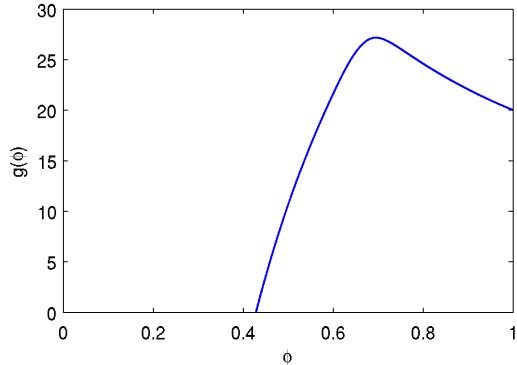
(a) Marginal Value of Capital
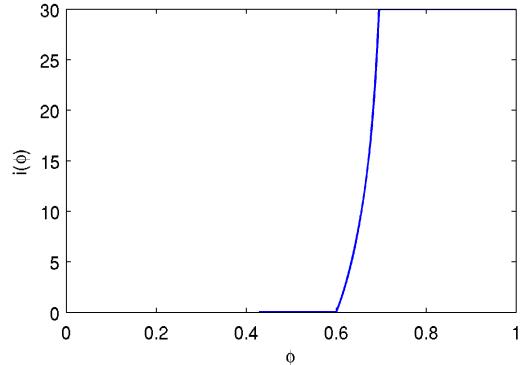
(b) Risky Assets to Total Capital

Figure 3.14: Case 8. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.06$, $\bar{i} = 30$ and $\hat{i} = 10^{-9}$.
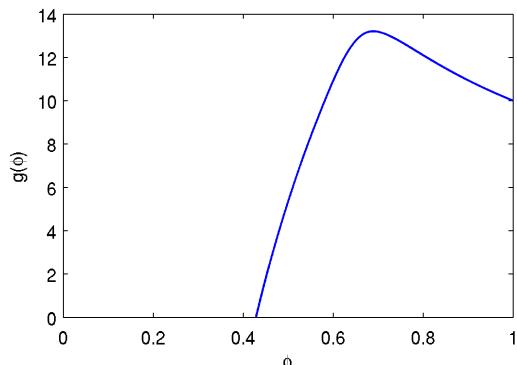
(a) Marginal Value of Capital
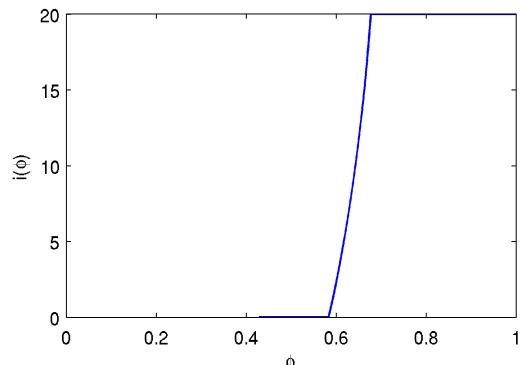
(b) Risky Assets to Total Capital

Figure 3.15: Case 9. Marginal value of capital for the good intermediary and demand for risky assets as share of total capital under management. Parameter values are: $r = 0.03$, $\mu_1 = 0.05$, $\mu_2 = 0.015$, $\sigma = 0.07$, $\rho = 0.07$, $\bar{i} = 20$ and $\hat{i} = 10^{-9}$.

References

Acharya, V., Davydenko, S. A., and Strebulaev, I. A. (2012). Cash holdings and credit risk. *Review of Financial Studies*, 25(12):3572–3609.

Acharya, V. V., Almeida, H., and Campello, M. (2007). Is cash negative debt? a hedging perspective on corporate financial policies. *Journal of Financial Intermediation*, 16(4):515–554.

Acharya, V. V. and Merrouche, O. (2012). Precautionary hoarding of liquidity and interbank markets: Evidence from the subprime crisis. *Review of Finance*, page rfs022.

Acharya, V. V. and Skeie, D. (2011). A model of liquidity hoarding and term premia in inter-bank markets. *Journal of Monetary Economics*, 58(5):436–447.

Ashcraft, A., McAndrews, J., and Skeie, D. (2011). Precautionary reserves and the interbank market. *Journal of Money, Credit and Banking*, 43(s2):311–348.

Banks, J. S. and Sobel, J. (1987). Equilibrium selection in signaling games. *Econometrica*, 55(3):647–661.

Basak, S. and Cuoco, D. (1998). An equilibrium model with restricted stock market participation. *Review of Financial Studies*, 11(2):309–341.

Beber, A., Brandt, M. W., and Kavajecz, K. A. (2009). Flight-to-quality or flight-to-liquidity? evidence from the euro-area bond market. *Review of Financial Studies*, 22(3):925–957.

Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6):1269–1295.

Bernanke, B. and Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review*, 79(1):14–31.

Bernanke, B., Gertler, M., and Gilchrist, S. (1996). The financial accelerator and the flight to quality. *The Review of Economics and Statistics*, pages 1–15.

Bernanke, B., Gertler, M., and Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics*, 1:1341–1393.

Board, S. and Meyer-ter Vehn, M. (2013). Reputation for quality. *Econometrica*, 81(6):2381–2462.

Bolton, P. and Harris, C. (1999). Strategic experimentation. *Econometrica*, 67(2):349–374.

Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *The American Economic Review*, 104(2):379–421.

Caballero, R. J. and Krishnamurthy, A. (2008). Collective risk management in a flight to quality episode. *The Journal of Finance*, 63(5):2195–2230.

Cuoco, D. and Kaniel, R. (2011). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics*, 101(2):264–296.

DeMarzo, P. M. and Sannikov, Y. (2006). Optimal security design and dynamic

capital structure in a continuous-time agency model. *The Journal of Finance*, 61(6):2681–2724.

Faingold, E. and Sannikov, Y. (2011). Reputation in continuous-time games. *Econometrica*, 79(3):773–876.

Friedman, M. and Schwartz, A. J. (1971). *A monetary history of the United States, 1867-1960*. Princeton University Press.

Fudenberg, D. and Levine, D. K. (1992). Maintaining a reputation when strategies are imperfectly observed. *The Review of Economic Studies*, 59(3):561–579.

Gale, D. and Yorulmazer, T. (2013). Liquidity hoarding. *Theoretical Economics*, 8(2):291–324.

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34.

Gertler, M. and Kiyotaki, N. (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of Monetary Economics*, 3(3):547–599.

Hansen, L. P. and Sargent, T. J. (2011). Robustness and ambiguity in continuous time. *Journal of Economic Theory*, 146(3):1195–1223.

He, Z. (2009). Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies*, 22(2):859–892.

He, Z. and Krishnamurthy, A. (2012). A model of capital and crises. *The Review of Economic Studies*, 79(2):735–777.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *The American Economic Review*, 103(2):732–70.

Holmstrom, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328.

Huang, C.-f. and Pages, H. (1992). Optimal consumption and portfolio policies with an infinite horizon: existence and convergence. *The Annals of Applied Probability*, pages 36–64.

Hugonnier, J. and Kaniel, R. (2010). Mutual fund portfolio choice in the presence of dynamic flows. *Mathematical Finance*, 20(2):187–227.

Kaniel, R. and Kondor, P. (2013). The delegated lucas tree. *Review of Financial Studies*, 26(4):929–984.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, volume 113. Springer Science & Business Media.

Liptser, R. and Shiryaev, A. N. (2001). *Statistics of random Processes: I. General Theory*. Springer Science & Business Media, 2nd edition edition.

Oksendal, B. (2013). *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media.

Panageas, S. and Westerfield, M. M. (2009). High-water marks: High risk appetites? convex compensation, long horizons, and portfolio choice. *The Journal of Finance*, 64(1):1–36.

Pástor, L. and Veronesi, P. (2009). Technological revolutions and stock prices. *The American Economic Review*, 99(4):1451–1483.

Revuz, D. and Yor, M. (2013). *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media.

Sannikov, Y. (2007). Games with imperfectly observable actions in continuous time. *Econometrica*, 75(5):1285–1329.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984.

Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *The Quarterly Journal of Economics*, pages 655–669.

Vayanos, D. (2004). Flight to quality, flight to liquidity, and the pricing of risk. Technical report, National Bureau of Economic Research.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies*, 12(5):975–1007.