

NBER WORKING PAPER SERIES

RISKY INVESTMENTS WITH LIMITED COMMITMENT

Thomas F. Cooley
Ramon Marimon
Vincenzo Quadrini

Working Paper 19594
<http://www.nber.org/papers/w19594>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2013

We would like to thank Youssef Cassis, Robert Lucas, Facundo Piguillem, Kim Schoenholtz, Danyang Xie and other seminar participants at Banca d'Italia, Banque de France, Bank of Turkey, Tel Aviv University, Hebrew University of Jerusalem, New York University, et al., for their comments on previous versions of this work. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w19594.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Thomas F. Cooley, Ramon Marimon, and Vincenzo Quadrini. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Risky Investments with Limited Commitment
Thomas F. Cooley, Ramon Marimon, and Vincenzo Quadrini
NBER Working Paper No. 19594
October 2013
JEL No. E25,E44,G01,G3

ABSTRACT

Over the last three decades there has been a dramatic increase in the size of the financial sector and in the compensation of financial executives. This increase has been associated with greater risk-taking and the use of more complex financial instruments. Parallel to this trend, the organizational structure of the financial sector has changed with the traditional partnership replaced by public companies. The organizational change has increased the competition for managerial talent, which may have weakened the commitment between investors and managers. We show how increased competition and the weaker commitment can raise the managerial incentives to undertake risky investment. In the general equilibrium, this change results in higher risk-taking, a larger and more productive financial sector with greater income inequality (within and across sectors), and a lower market valuation of financial institutions.

Thomas F. Cooley
Department of Economics
Stern School of Business
44 West 4th Street, Room 7-88
New York, NY 10012-1126
and NBER
tcooley@stern.nyu.edu

Vincenzo Quadrini
Department of Finance and Business Economics
Marshall School of Business
University of Southern California
701 Exposition Boulevard
Los Angeles, CA 90089
quadrini@usc.edu

Ramon Marimon
European University Institute
Department of Economics
Via della Piazzuola 43
50133 Florence
ITALY
and UPF - BarcelonaGSE
and also NBER
ramon.marimon@eui.eu

1 Introduction

The past several decades have been characterized by dramatic changes in the size and structure of financial firms in the United States and elsewhere. What was once an industry dominated by partnerships has evolved into a much more concentrated sector dominated by large public firms. In this paper we argue that this evolution has altered the structure of contractual arrangements between investors and managers in ways that weakened commitment and increased the managers' incentives to undertake risky investments. At the aggregate level, the change resulted in a larger financial sector and greater income inequality.

The increase in the size and importance of the financial sector in the United States has been documented by Phillipon (2008) and Phillipon and Resheff (2009).¹ This is also shown in Figure 1 which plots the shares of the financial industry in value added and employment since 1970. The contribution of the finance industry to GDP doubled in size between 1970 and 2011. The share of employment has also increased but by less than the contribution to value added. This is especially noticeable starting in the mid 1980s when the share of employment stopped growing while the share of value added continued to expand. Accordingly, we observe a significant increase in productivity compared to the remaining sectors of the economy.

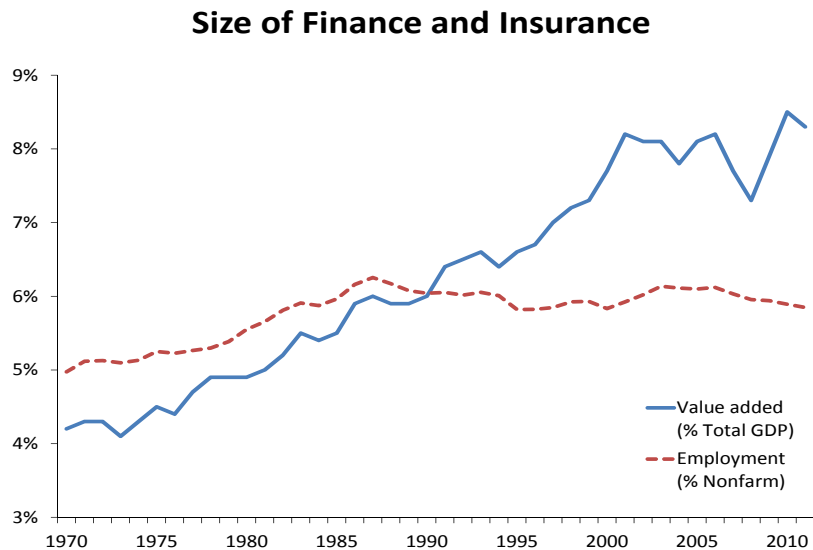


Figure 1: Share of Value Added and Employment

The increase in size was also associated with a sharp increase in compensation. Clementi and Cooley (2009) show that between 1993 and 2006 the average compensation levels of CEOs in the financial sector increased from parity with other sectors of

¹Phillipon and Resheff (2013) also study the growth of finance in a cross section of countries and find that the share of the financial sector in many of them is significantly higher today than at any point in the past century.

the economy to nearly double. At the same time compensation of managers became more unequal in the financial sector. Figure 2 shows the evolution of the income share of the top 5% of managerial positions in the sector compared to other occupations.

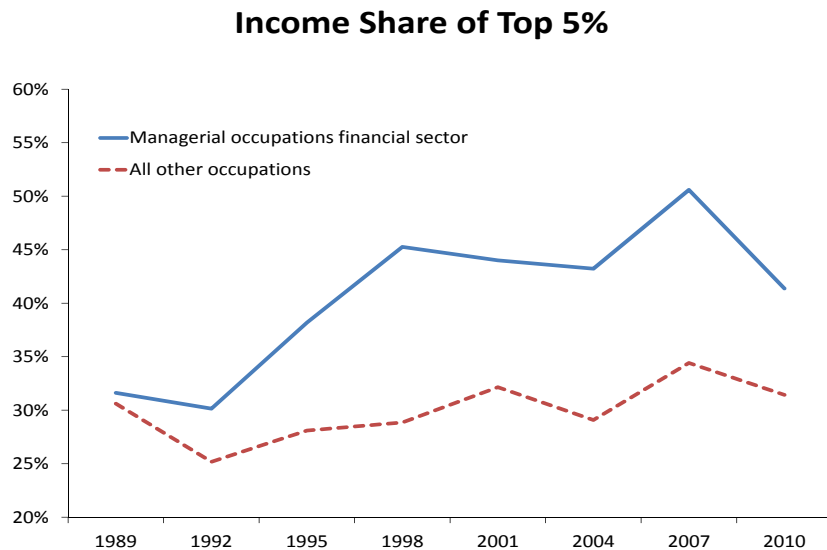


Figure 2: Share of the top 5%

This period of increased size and importance of the financial sector followed significant changes in the organizational form of financial firms—especially the transformation from partnerships to public corporations—that had two important effects. The first was to increase competition in the financial sector raising the demand for managers. The second was to alter the structure of contractual arrangements between investors and managers in ways that weakened commitments. As we will see, the combination of these two effects increased the managers’ incentives to undertake risky investments and generated greater income inequality within and between sectors.

Historically, it was common for investment firms to be organized as partnerships. Many argued that this was a preferred form of organization because in a partnership, managers and investors were the same people and it was the partners own assets that were at risk when risky investments were taken. Effectively, the separation between ownership and investment control is minimized, reducing the agency issues. Public companies, on the other hand, are organizational structures with significant separation between ownership (shareholders) and investment control (managers), and it is well understood that this organizational form is characterized by significant agency issues.²

Until 1970 the New York Stock Exchange prohibited member firms from being public companies. When the organizational restriction on financial companies was relaxed,

²This is largely consistent with the literature on incomplete contract theory. According to Grossman and Hart (1986) and Hart and Moore (1990), more efficient organizational forms are those where the agents who control the investment surplus own a larger share of the assets.

there was a movement to go public and partnerships began to disappear. Merrill Lynch went public in 1971, followed by Bear Stearns in 1985, Morgan Stanley in 1985, Lehman Brothers in 1994 and Goldman Sachs in 1999. Other venerable investment banks were taken public and either absorbed by commercial banks or converted to bank holding companies. Today, there are very few partnerships remaining and they are small. The same evolution occurred in Britain where the closed ownership Merchant Banks virtually disappeared.³

The partnership form and its customs had some important implications for managerial mobility. The capital in a partnership and the ownership shares are typically relatively illiquid so it was difficult for partners to liquidate their ownership positions and move to other firms. Also important is the process of becoming a partner. In the typical firm, new professionals are hired as associates and, after a trial period, they are either chosen to be partners or released. In this environment separation is viewed as a signal of inferior performance, thus affecting the external option of a financial professional. Becoming a partner, on the other hand, represented a firm commitment to continued employment on the part of the other partners. The transition to corporations, however, changed significantly the nature of contracts and competition in the financial sector.⁴

As the structure of financial firms changed, so did the market valuation of them. The market does not seem to value highly large complex financial institutions. Figure 3 shows the evolution of the ratio of average market value of equity to book value of equity for publicly listed financial and nonfinancial firms since 1970 and shows that, starting in the early 1980's, the market valuation of financial firms has been flat while for nonfinancial firms it has continued to grow. The fact that the market values the financial sector relatively less, compared to the rest of the economy, may be a reflection of compensation practices in firms where managers retain so much of the surplus.⁵

In sum, we have emphasized four changes that are associated with the evolution of the financial sector from a partnership form of organization to public corporations: (*i*) greater risk-taking; (*ii*) a larger and more productive financial sector; (*iii*) lower stock

³The fact that member firms were allowed to become public companies does not tell us why they chose to do so. In several cases firms were simply acquired by public companies but in others it was an important strategic decision. Charles Ellis (2008) in his history of Goldman Sachs—the last major firm to go public—suggests that the major motive for financial partnerships to become public was to increase capital for their proprietary trading operations through an IPO. The transition from partnership to public corporation also had implication for liability of managers. We don't explicitly address this aspect of the organizational change.

⁴Roy Smith, a former partner at Goldman Sachs described the evolution of the relationship between compensation and firm structure as follows: "In time there was an erosion of the simple principles of the partnership days. Compensation for top managers followed the trend into excess set by other public companies. Competition for talent made recruitment and retention more difficult and thus tilted negotiating power further in favor of stars. You had to pay everyone well because you never knew what next year would bring, and because there was always someone trying to poach your best trained people, whom you didn't want to lose even if they were not superstars. Consequently, bonuses in general became more automatic and less tied to superior performance. Compensation became the industry's largest expense, accounting for about 50% of net revenues", *Wall Street Journal* February 7, 2009.

⁵Since the financial crisis, compensation in the securities industry has increased by 8.7% annually. Currently nearly half of all revenues are earmarked for compensation and it has been higher in the past.

Market to Book Value of Assets

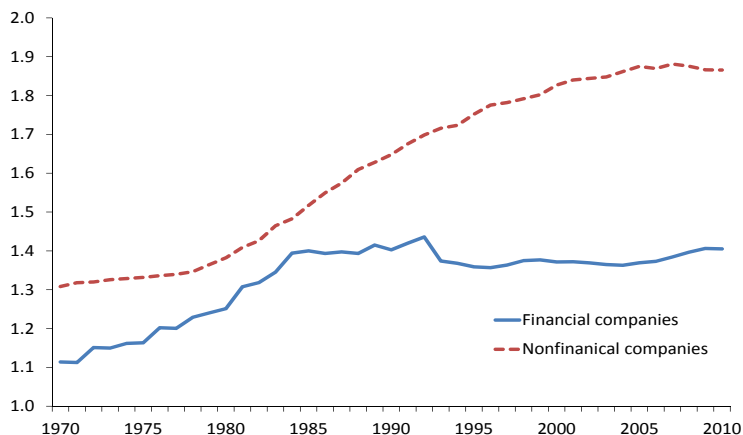


Figure 3: Average Market Value of Equity/Book Values of Equity

market evaluation of financial institutions, and (*iv*) greater income inequality (within and in relation to other sectors). Existing models capture some of these facts but not all of them simultaneously. We are also not aware of any study that connects change in the organizational structure of the financial sector with the increased competition for managerial talent.⁶ In this study, instead, we propose a model that generates all of the above facts as a consequence of the organizational change that has taken place in the financial sector during the last three decades.

We study a model where investors compete for and hire managers to run investment projects, with each investor-manager pair representing a financial firm. A key feature of the model is that production depends on the human capital of the manager which can be enhanced, within the firm, with costly investment. Human capital accumulation can be understood as acquiring new skills by engaging in risky financial innovations (e.g. implementing new financial instruments which may or may not have positive returns). Since part of the accumulated human capital can be transferred outside the firm by the manager, there is a conflict of interest between the investor and the manager.

In this environment, the investment desired by the investor may be smaller than the investment desired by the manager because the cost is incurred by the firm while the benefits are shared. This implies that, if the investor cannot control the investment policy either directly or indirectly through a credible compensation scheme, the manager has an

⁶For example, Cheng, Hong and Scheinkman (2012) and, in a general equilibrium framework, Edmans and Gabaix (2011) explain how in a Principal-Agent relationship, with a fixed sharing rule, an *exogenous* increase in risk can result in higher compensation, since risk-averse financial managers must satisfy their participation and incentive constraints. Bolton, Santos and Scheinkman (2012) argue that it is “cream skimming” in the more opaque financial transactions—those taking place in over-the-counter or bespoke markets—that have encouraged excessive compensation of financial managers and the excessively large share of GDP of the financial services industry.

incentive to deviate from the optimal policy simply because she does not internalize the full cost of the investment. The goal of the paper is to characterize the investment and compensation policies that result from the (constrained) optimal contract and show how these policies change when the competition for managers increases and the enforcement of contracts weakens.

The basic framework that is often used to study executive compensation is adapted from the principle-agent model of dynamic moral hazard by Spear and Srivastava (1987).⁷ An assumption typically made in this class of models is that the outside option of the agent is exogenous. As argued above, however, an important consequence of the demise of the partnership form is that financial managers are no longer constrained by the limited liquidity of the portion of their wealth that is tied to the firm and it is easier for them to seek outside employment. Since the value of seeking outside employment depends on the market conditions for managers, it becomes important to derive these conditions endogenously in general equilibrium.

A second assumption typically made in principal-agent models is that investors fully commit to the contract. However, the clearer separation between investors and managers that followed the transformation of financial partnerships to public companies and the associated “competition for managerial talent”, could have also reduced the commitment of investors. Therefore, in this paper we relax both assumptions: we endogenize the outside option of managers which will be determined in general equilibrium and we allow for the limited commitment of investors.⁸

A main result of this paper is that, whether an increase in competition for managerial talent results in higher (or lower) risk-taking depends on whether there is double-sided (or one-sided) limited commitment. When the investor can commit, the constrained-efficient contract takes into account that in a more competitive environment the risk-averse manager has a greater incentive to take risks in order to enhance her outside value, but does so at the expense of the firm.⁹ Instead, when the investor cannot fully commit to future compensation, it is in the manager’s interest to choose the investment level that maximizes her outside value and this increases with competition.

⁷Among the models in this class see, for example, Wang (1997), Quadrini (2004), Clementi and Hopenhayn (2006), Fishman and DeMarzo (2007). Albuquerque and Hopenhayn (2004) can also be considered within this class of models although the frictions are based on limited enforcement rather than information asymmetry.

⁸Although in a different set-up, Cooley, Marimon and Quadrini (2004) endogenized the outside value of entrepreneurs but kept the assumption that investors commit to the long-term contract. Marimon and Quadrini (2011) relaxed both assumptions and, using a model without uncertainty, showed that differences in “barriers to competition”, can result in income differences across countries. In these two papers, however, uncertainty does not play a significant role while it is central to the analysis of the current paper.

⁹Notice that, as Cheng *et al.* (2012) have shown, *exogenous* higher risk typically results in higher compensation for managers (see Footnote 6). Consistent with this, we show that when the level of risk is an *endogenous* variable, the constrained-efficient contract recommends lower risk. In fact, within the standard one-sided limited commitment Principal-Agent problem, Edmans and Gabaix (2011) and Bolton *et al.* (2012) obtain changes in aggregate risk as a result of the interplay between the principal-agent problem and an assignment/reallocation problem. We obtain the increase in risk from *competition* and *weak commitment*.

To make the outside value of managers endogenous and to study the implications for the whole economy, we embed the micro structure in a general equilibrium model with two sectors—financial and nonfinancial. With this general framework we study the consequences of the organizational changes which, as discussed above, had two effects: it increased competition in the financial sector raising the demand for managers, and it altered the structure of contractual arrangements between investors and managers in ways that weakened commitments. These two effects are formalized in the model by a lower cost to create jobs in the financial sector and by a shift to a regime where investors do not commit to the contract (double-sided limited commitment). We then show that these structural changes can generate (i) greater risk-taking; (ii) larger share (and higher relative productivity) of the financial sector; (iii) lower stock market valuation of financial institutions; (iv) greater income inequality within and between sectors.

The organization of the paper is as follows. In Section 2 we describe the environment and characterize the optimal contract under different assumptions about commitment. Section 3 embeds the micro structure in a general equilibrium model. Section 4.1 provides a numerical characterization of the equilibrium and relates its properties to the empirical facts that motivate the paper. Section 5 concludes.

2 The model

We start with the description of the financial sector and the contracting relationships that are at the core of the model. After the characterization of the financial sector, we will embed it in a general equilibrium in Section 3.

The financial sector is characterized by firms regulated by a contract between an investor, the owner of the firm, and a manager. We should think of managers as skilled workers who have the ability to run the firm and implement innovative projects.

Managers are characterized by their human capital h_t and are endowed with one unit of time that can be used in two alternative activities: production and innovation. Denote by λ_t the time allocated to innovating in period t . Then the output produced by the firm in period $t + 1$ is equal to

$$Y_{t+1} = y(\lambda_t)h_t, \tag{1}$$

where the function $y(\cdot)$ satisfies $y' < 0$, $y'' > 0$, $y(1) = 0$. Therefore, output increases with the manager's human capital, h_t , and decreases with the time allocated to innovation, λ_t (since the manager allocates less time managing production). The convexity assumption captures the idea that, as the manager spends less time producing, the ordinary operation of the firm becomes less efficient. Notice that production activities performed in period t generate output in period $t + 1$. The significance of this assumption will be emphasized below.

Innovation activities consist of the development of a new implementable project or idea of size i_{t+1} according to the technology

$$i_{t+1} = h_t \lambda_t \varepsilon_{t+1},$$

where λ_t is the manager's time allocated to innovation activities and $\varepsilon_{t+1} \in \{0, \bar{\varepsilon}\}$ is an i.i.d. stochastic variable that takes the value of zero with probability $1 - p$ and 1 with probability p .

We think of λ_t as the investment to generate a new implementable project i_{t+1} whose outcome is uncertain because of the stochastic variable ε_{t+1} . A feature of the innovation technology is that the standard deviation of i_{t+1} is linear in h_t and λ_t . Higher values of λ_t are associated with greater uncertainty and, therefore, higher risk.

If implemented, the new project enhances the human capital of the manager according to $h_{t+1} = h_t + i_{t+1}$. However, for the manager's human capital to grow, it is essential that the new project is implemented. This would happen if the manager continues to be employed in a financial firm. If the new project is not implemented—for instance, if the manager leaves the financial sector—her human capital remains h_t . Therefore, if new projects are implemented after their development stage, they become *embedded* human capital. Otherwise they fully depreciate. The importance of this assumption will become clear later.¹⁰

To use a compact notation, we define $g(\lambda_t, \varepsilon_{t+1}) = 1 + \lambda_t \varepsilon_{t+1}$ the gross growth rate of human capital, provided the manager remains employed. Then, the evolution of human capital can be written as

$$h_{t+1} = g(\lambda_t, \varepsilon_{t+1})h_t. \quad (2)$$

Investors are risk-neutral and they are the residual claimants to the output produced by the firm. Their expected lifetime utility is

$$V_0 = E_t \sum_{t=0}^{\infty} \beta^t (\beta Y_{t+1} - C_t).$$

The expected lifetime utility of managers takes the form

$$Q_0 = E_t \sum_{t=0}^{\infty} \beta^t [u(C_t) - e(\lambda_t)],$$

where C_t is consumption and $e(\lambda_t)$ is the dis-utility from innovation activities. The period utility satisfies $u' > 0$, $u'' < 0$ and $e' > 0$, $e'' > 0$, $e(0) = 0$ and $e(1) = \infty$.

There are two types of cost associated with financial innovation. The first is the loss of production as the manager spends more time innovating. The second is the manager's dis-utility from innovating. A key difference between these two costs is that the first is incurred by the firm while the second is incurred by the manager. This creates a wedge between who pays the cost of the innovation and who enjoys the benefits: If the manager chooses to quit, the production cost is incurred by the firm but the benefits go to the manager in the form of increased human capital (provided that the manager finds

¹⁰The assumption that newly developed projects depreciate if not implemented while the pre-existing human capital does not depreciate is not essential for the qualitative properties of the model. It is only made to maintain the linear homogeneity in h_t . This property does not hold if we assume that the whole human capital depreciates (old and new) when the manager moves away from the financial sector.

occupation in another financial firm). This asymmetry plays an important role for the results of the paper.

Managers have the option to quit and search for an offer from a new firm. If they choose to quit, they will receive an offer with probability $\rho \in [0, 1]$. The probability ρ captures the degree of *competition* for managers, that is, the ease with which a manager finds occupation in the financial sector after quitting the firm. Higher values of ρ denote a more competitive financial sector. Since we are assuming that an implementable project of size i_{t+1} fully depreciates if not implemented in a firm, the human capital of a manager who chooses to quit at the beginning of period $t + 1$ will be $h_t + i_{t+1}$ only if she receives an offer. Otherwise, the human capital remains h_t .

Denote by $\underline{Q}_{t+1}(h_t)$ the outside value at the beginning of period $t + 1$ without an external offer and by $\overline{Q}_{t+1}(h_{t+1})$ the outside value with an offer. The expected outside value at $t + 1$ of a manager with previous human capital h_t is equal to

$$D(h_t, h_{t+1}, \rho) = (1 - \rho) \cdot \underline{Q}_{t+1}(h_t) + \rho \cdot \overline{Q}_{t+1}(h_{t+1}), \quad (3)$$

where $h_{t+1} = h_t(1 + \lambda_t \varepsilon_{t+1})$

For the moment we take ρ , $\underline{Q}_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ as given. At this stage we only assume that $\underline{Q}_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ are strictly increasing and differentiable, which implies $D_{2,3} > 0$. However, when we extend the model to a general equilibrium in Section 3, the probability of an external offer and the outside values with and without an offer will be derived endogenously. This is an important innovation of our model and will be central for some of the results.

In addition to having the ability to quit, the manager has full control over the choice of λ_t . Full control is allowed by the assumption that λ_t is directly observable only by the manager. The investor can only infer the actual value of λ_t in the next period after the realization of output Y_{t+1} . This implies that, in absence of proper incentives, the λ_t chosen by the manager may not maximize the surplus of the partnership. Therefore, there are two sources of frictions on the side of the manager: the ability to quit and the discretion to choose any value of λ_t .

Definition 1 *A **contract** between an investor and a manager with initial human capital h_0 consists of sequences of payments to the manager $\{C(\mathbf{H}^t, \mathbf{\Lambda}^t)\}_{t=0}^{\infty}$ and investments $\{\lambda(\mathbf{H}^t, \mathbf{\Lambda}^t)\}_{t=0}^{\infty}$, conditional on the observed history of human capital $\mathbf{H}^t = (h_0, \dots, h_t)$ and investment $\mathbf{\Lambda}^t \equiv (\lambda_0, \dots, \lambda_{t-1})$.*

Notice that the payment made to the manager in period t is not conditional on the innovation λ_t chosen by the manager in period t . This is because λ_t becomes public information (by observing production) only in the next period.

2.1 The optimal contract with one-sided limited commitment

We first characterize the optimal contract when the investor commits but the manager does not (one-sided limited commitment). In this environment the manager could quit the firm at any point in time and could choose any investment λ_t . The optimal contract

can be characterized by solving a planner's problem that maximizes the weighted sum of utilities for the investor and the manager but subject to a set of constraints. These constraints guarantee that the allocation chosen by the planner is enforceable in the sense that both parties choose to participate and the manager has no incentive to take actions other than those prescribed by the contract. We first characterize the key constraints and then we specify the optimization problem.

The allocation chosen by the planner must be such that the value of the contract for the manager is not smaller than the value of quitting. This gives rise to the *enforcement constraint*

$$E_{t+1} \sum_{n=0}^{\infty} \beta^n \left[u(C_{t+1+n}) - e(\lambda_{t+1+n}) \right] \geq D(h_t, h_{t+1}, \rho), \quad t \geq 0. \quad (4)$$

A second constraint takes into account that the manager has full control of the investment λ_t . The manager could deviate from the λ_t recommended by the planner since, through the choice of λ_t , she can affect the outside value. Thus, the allocation must satisfy an incentive-compatibility constraint insuring that the manager does not deviate from the recommended investment policy.

Denote by $\hat{\lambda}_t$ the investment chosen by the manager when she deviates from the recommended λ_t . This maximizes the outside value net of the cost of effort, that is,

$$\hat{\lambda}_t = \arg \max_{\lambda \in [0,1]} \left\{ -e(\lambda) + \beta E_t D(h_t, g(\lambda, \varepsilon_{t+1})h_t, \rho) \right\}. \quad (5)$$

Since the outside value of the manager is differentiable, the optimal deviation solves the first-order condition

$$e_\lambda(\hat{\lambda}_t) \geq \beta E_t D_2(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho) g_\lambda(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \quad (6)$$

which is satisfied with equality if $\hat{\lambda}_t > 0$.

We can now see the importance of the assumption that the manager faces the effort dis-utility from innovating. In absence of this, the optimal deviation $\hat{\lambda}_t$ would be 1. Instead, with $e_\lambda(1) = \infty$, the optimal deviation is interior in the interval $[0, 1]$ and is affected by a change in the outside value.

Given the optimal deviation $\hat{\lambda}_t$, the *incentive-compatibility constraint* at t is

$$-e(\lambda_t) + \beta E_t \sum_{n=0}^{\infty} \beta^n \left(u(C_{t+n+1}) - e(\lambda_{t+n+1}) \right) \geq -e(\hat{\lambda}_t) + \beta E_t D(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho). \quad (7)$$

Notice that C_t does not appear in (7) because current consumption cannot be contingent on current investment, λ_t , because the investment becomes public information in the next period. This limits the ability of the planner to punish the manager by cutting consumption in the current period. The manager can be punished in the next period, once the investment becomes public information. At this stage, however, the manager has the option of quitting which imposes a lower bound on the possible punishment.

We now have all the ingredients to write down the optimization problem solved by the planner in a regime with one-sided limited commitment. Let $\tilde{\mu}_0$ be the planner's weight assigned to the manager and normalize to 1 the weight assigned to the investor. We can then write the planner's problem as

$$\begin{aligned} \max_{\{C_t, \lambda_t\}_{t=0}^{\infty}} \quad & E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left(\beta y(\lambda_t) h_t - C_t \right) + \tilde{\mu}_0 \sum_{t=0}^{\infty} \beta^t u(C_t) \right\} \\ \text{s.t.} \quad & (2), (4), (7). \end{aligned} \quad (8)$$

The problem is also subject to initial participation constraints for both the investor and the manager which, for simplicity, we have omitted. They only restrict the admissible values for the weight $\tilde{\mu}_0$.

Following Marcet and Marimon (2011), the problem can be written recursively as

$$\begin{aligned} \widetilde{W}(h, \tilde{\mu}) = \min_{\tilde{\chi}, \tilde{\gamma}(\varepsilon')} \max_{C, \lambda} \quad & \left\{ \beta y(\lambda) h - C + \tilde{\mu} \left(u(C) - e(\lambda) \right) - \tilde{\chi} \left(e(\lambda) - e(\hat{\lambda}) \right) + \right. \\ & \left. \beta E \left[\widetilde{W}(h', \tilde{\mu}') - \tilde{\chi} D \left(h, g(\hat{\lambda}, \varepsilon') h, \rho \right) - \tilde{\gamma}(\varepsilon') D(h, h', \rho) \right] \right\} \\ \text{s.t.} \quad & h' = g(\lambda, \varepsilon') h, \quad \tilde{\mu}' = \tilde{\mu} + \tilde{\chi} + \tilde{\gamma}(\varepsilon'), \end{aligned} \quad (9)$$

where $\tilde{\gamma}(\varepsilon')$ is the Lagrange multiplier for the enforcement constraint (4) and $\tilde{\chi}$ is the Lagrange multiplier for the incentive-compatibility constraint (7).

The function $\widetilde{W}(h, \tilde{\mu})$ is related to the value of the contract for the investor, $\widetilde{V}(h, \tilde{\mu})$, and to the value for the manager, $\widetilde{Q}(h, \tilde{\mu})$, by the equation $\widetilde{W}(h, \tilde{\mu}) = \widetilde{V}(h, \tilde{\mu}) + \tilde{\mu} \widetilde{Q}(h, \tilde{\mu})$. An environment with full commitment is just a special case with $\tilde{\gamma}(\varepsilon') = \tilde{\chi} = 0$. Another special case is when λ_t is controlled by the investor, in which case we impose $\tilde{\chi} = 0$.

Differentiating problem (9) with respect to C we obtain the optimality condition

$$C_t = u_c^{-1} \left(\frac{1}{\tilde{\mu}_t} \right), \quad (10)$$

which characterizes the **consumption policy** as a function of the state variable $\tilde{\mu}_t$. This variable evolves according to the law of motion $\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})$.

It is useful to consider the *normalized* manager's weight, or manager's share of the surplus, $\mu_t = \tilde{\mu}_t/h_t$. Since, income is proportional to h_t (recall (1)), the manager's share plays a key role in our analysis. The normalized value functions take the form $W(h, \mu) \equiv \widetilde{W}(h, h\mu)/h = V(h, \mu) + \mu Q(h, \mu)$, where $Q(h, \mu) \equiv \widetilde{Q}(h, h\mu)$ and $V(h, \mu) \equiv \widetilde{V}(h, h\mu)/h$. Furthermore, since the technology and the preferences of the investor are linear in h , we have that $V(h, \mu) = v(\mu)$. therefore, $W(h, \mu) = v(\mu) + \mu Q(h, \mu)$. Similarly, we can normalize the multipliers— $\chi_t = \tilde{\chi}_t/h_t$ and $\gamma_t(\varepsilon_{t+1}) = \tilde{\gamma}_t(\varepsilon_{t+1})/h_t$ —to obtain the law of motion of the manager's share $\mu_{t+1} = (\mu_t + \chi_t + \gamma_t(\varepsilon_{t+1}))/g(\lambda_t, \varepsilon_{t+1})$.

If contracts were perfectly enforceable, χ_t and $\gamma_t(\varepsilon_{t+1})$ would be both equal to zero, but then if $\lambda_t > 0$, $E\mu_{t+1} < \mu_t$, since $Eh_{t+1} = (1+p\lambda_t)h_t$. Given that optimal investments satisfy $\lambda_t > 0$, the normalized share of the surplus μ_t converges to zero (i.e. it is a lower-bounded submartingale). However, with one-sided limited commitment, the enforcement and incentive-compatibility constraints set a lower bound on μ_t . Thus, with one-sided limited commitment the manager has a minimum share guaranteed. A downside of this is that the contract does not provide full insurance since the manager's consumption increases stochastically.

The *investment policy* is characterized by the first-order condition with respect to λ . Using $g(\lambda_t, \varepsilon_{t+1}) = 1 + \lambda_t \varepsilon_{t+1}$ and the above normalization, the optimality condition can be written as

$$(\mu_t + \chi_t) e_\lambda(\lambda_t) - \beta y_\lambda(\lambda_t) \geq \beta E_t \left[\left(v(\mu_{t+1}) + (\mu_t + \chi_t + \gamma_t(\varepsilon_{t+1})) Q_h(h_{t+1}, \mu_{t+1}) - \gamma_t(\varepsilon_{t+1}) D_2(h_t, h_{t+1}, \rho) \right) \varepsilon_{t+1} \right], \quad (11)$$

which is satisfied with equality when $\lambda_t > 0$. See Appendix B for more details.

The left-hand side of (11) is the marginal (direct) cost of investment per unit of human capital, which is increasing in λ_t , μ_t and χ_t . The right-hand-side is the expected marginal benefit from investing, net of participation costs. With **full commitment** the marginal (direct) cost decreases, since $\chi_t = 0$, and μ_t converges to zero. Notice that, with full commitment, there are no inter-temporal participation costs (i.e. $\gamma(\varepsilon_{t+1}) = 0$). Furthermore, $Q_h(h_{t+1}, \mu_{t+1}) = 0$, since the manager's consumption is constant (only depends on $\tilde{\mu}_0$) and the investment effort is independent of h . Therefore, only $v(\mu_{t+1})$ matters and, given that $v'(\mu) < 0$, the right-hand side of (11) increases. In sum, λ_t^* is increasing.

With **one-sided limited commitment**, a binding enforcement constraint is an additional cost to investment: the cost of having to compensate the manager to keep her in the firm, whenever $\gamma(\varepsilon_{t+1}) > 0$. This lowers the increase in the right-hand side of (11). Furthermore, in relation to the left-hand side of (11), limited enforcement and/or incentive compatibility constraints raise a manager's marginal effort-cost when they are binding and prevent μ from converging to zero. In sum, given (h, μ) , the investment policy with one-sided limited commitment is less risky when the enforcement and incentive constraints are binding.

We are particularly interested in the effect of increasing competition (higher ρ) on the optimal investment policy. In an economy with higher ρ managers have better outside opportunities implying that the initial μ_0 is higher, for a given h_0 . This redistribution affects investment. For instance, with **full commitment**, it affects the λ_0^* : it increases the right-hand side of (11) and lowers $v(\mu_0)$. Therefore, a contract that starts in a higher ρ economy it is characterized by a higher μ_0 and a lower expected risk profile. But, once the investor and the manager are committed to the contract increasing competition has no effect on investment.

With **one-sided limited commitment** an economy with a relative high ρ is also characterised by a higher μ_0 and a lower expected risk profile. However, increasing

competition after the contract starts has an effect on investment. There is a direct effect on the participation cost when the enforcement constraint is (or becomes) binding, through $D_{2,3}$ in the right and-hand side of (11) and an indirect effect through the increases in the multipliers of the enforcement and incentive constraints, $\gamma(\varepsilon')$ and χ . In sum, we obtain the following result (formally proved in Appendix B).

Proposition 1 *Suppose that, at (h, μ) , the optimal investment is interior, that is $\lambda^* \in (0, 1)$. If $Q_{h,\mu}(h', \mu') \leq 0$ then more competition for managers (higher ρ) results in a lower optimal investment λ^* when the enforcement constraint is, or becomes, binding.*

As we discuss in Appendix B, the assumption $Q_{h,\mu}(h', \mu') \leq 0$ is fairly general and, in particular, it is satisfied if the manager has CRRA preferences for consumption with an intertemporal elasticity of substitution less than or equal to one. In Subsection 2.3 we present in more detail the *log* case of unitary elasticity of substitution, where $Q_{h,\mu}(h', \mu') = 0$, which is then used in the rest of the paper.

2.2 The optimal contract with double-sided limited commitment

The law of motion $\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})$ captures the investor's commitment to fulfill promises made to the manager. With double-sided limited commitment the investor does not commit to fulfill his promises and renegotiates whenever the value of the contract for the manager exceeds the outside value. This implies that the value of $\tilde{\mu}_t$ chosen in the previous period becomes irrelevant for the new $\tilde{\mu}_{t+1}$ chosen in the current period. Under these conditions, the manager has the incentive to choose the investment that maximizes the outside value as defined in (5), that is, $\lambda_t = \hat{\lambda}_t$. Thus, the incentive-compatibility constraint and the multiplier $\tilde{\chi}_t$ become irrelevant, and the optimal contract solves

$$W(h, \tilde{\mu}) = \min_{\tilde{\gamma}(\varepsilon')} \max_C \left\{ \beta y(\hat{\lambda})h - C + \tilde{\mu} \left(u(C) - e(\hat{\lambda}) \right) + \beta E \left[W \left(g(\hat{\lambda}, \varepsilon')h, \tilde{\mu}' \right) - \tilde{\gamma}(\varepsilon') D \left(h, g(\hat{\lambda}, \varepsilon')h, \rho \right) \right] \right\} \quad (12)$$

$$\text{s.t.} \quad \tilde{\mu}' = \tilde{\gamma}(\varepsilon').$$

The contract with double-sided limited commitment simply prescribes a consumption plan which is determined by (10) with $\tilde{\mu}' = \tilde{\gamma}(\varepsilon')$, and the investment is $\hat{\lambda}$, which is the solution to (6). Since $D_{2,3} > 0$, an increase in competition captured by the parameter ρ increases the right-hand-side of (6), that is, it increases the marginal benefit of investing more for the manager. This is stated formally in the next proposition.

Proposition 2 *Consider the environment with double-sided limited commitment and suppose that $\hat{\lambda} \in (0, 1)$. Then a higher ρ results in a higher investment $\hat{\lambda}$.*

Propositions 1 and 2 show that the impact of higher competition on risk-taking depends crucially on whether both agents can commit to the contract. We should expect increasing competition to result in increased risk-taking only when there is limited commitment from both investors and managers.

2.3 The normalized contract

Since human capital grows on average over time, so does the value of the contract for both the manager and the investor. It is convenient to normalize the growing variables so that we can work with stationary variables. The normalization will be especially convenient in Section 3 when we embed the financial sector in a general equilibrium. This will be facilitated by specifying the utility of managers in log-form and by assuming that the outside values of managers—which at this stage are exogenous—take special functional forms.

Assumption 1 *The utility function and the outside values of managers take the forms*

$$\begin{aligned} u(C) - e(\lambda) &= \ln(C) + \alpha \ln(1 - \lambda), \\ \underline{Q}_{t+1}(h_t) &= \underline{q} + \mathcal{B} \ln(h_t), \\ \overline{Q}_{t+1}(h_{t+1}) &= \overline{q} + \mathcal{B} \ln(h_{t+1}), \end{aligned}$$

where \underline{q} , \overline{q} and $\mathcal{B} \equiv \frac{1}{1-\beta}$ are constant.

Although the functional forms for the outside values may seem arbitrary at this stage, we will see that in the extension to the general equilibrium they take these forms. Also notice that Assumption 1 guarantees that these functions are differentiable and strictly increasing as we assumed earlier, which in turn implies $D_{2,3} > 0$.

We are now ready to normalize all growing variables. The value of the contract for the investor can be expressed recursively as $V_t = \beta y(\lambda_t) h_t - C_t + \beta E_t V_{t+1}$ and can be normalized to

$$v_t = \beta y(\lambda_t) - c_t + \beta E_t g(\lambda_t, \varepsilon_{t+1}) v_{t+1}, \quad (13)$$

where $v_t = V_t/h_t$ and $c_t = C_t/h_t$.

The value of the contract for a manager can be expressed recursively as $Q_t = \ln(C_t) + \alpha \ln(1 - \lambda_t) + \beta E_t Q_{t+1}$. If we subtract $\mathcal{B} \ln(h_t)$ on both sides and add and subtract $\beta \mathcal{B} E_t \ln(h_{t+1})$ on the right hand side, we obtain

$$Q_t - \mathcal{B} \ln(h_t) = \ln(c_t) + \alpha \ln(1 - \lambda_t) + \beta \mathcal{B} E_t \ln\left(\frac{h_{t+1}}{h_t}\right) + \beta E_t \left[Q_{t+1} - \mathcal{B} \ln(h_{t+1}) \right].$$

Defining $q_t = Q_t - \mathcal{B} \ln(h_t)$, we can rewrite the above expression more compactly as

$$q_t = \ln(c_t) + \alpha \ln(1 - \lambda_t) + \beta E_t \left[\mathcal{B} \ln\left(g(\lambda_t, \varepsilon_{t+1})\right) + q_{t+1} \right]. \quad (14)$$

This provides the normalized value of the manager in recursive form.

The enforcement constraint for the manager after the realization ε_{t+1} is

$$Q_{t+1}(h_{t+1}) \geq (1 - \rho) \cdot \underline{Q}_{t+1}(h_t) + \rho \cdot \overline{Q}_{t+1}(h_{t+1}).$$

Using $q_{t+1} = Q_{t+1}(h_{t+1}) - \mathcal{B} \ln(h_{t+1})$ and the functional forms specified in Assumption 1, the enforcement constraint (7) can be written as

$$q_{t+1} \geq (1 - \rho)\underline{q} + \rho\bar{q} - (1 - \rho)\mathcal{B} \ln\left(g(\lambda_t, \varepsilon_{t+1})\right). \quad (15)$$

The right-hand-side depends on λ_t (provided that $\rho < 1$). Thus, investment affects the outside value of the manager and, when the enforcement constraint is binding, it affects compensation. This property is a direct consequence of the assumption that the outside value of the manager *without* an external offer depends on h_t , while the outside value *with* an external offer depends on h_{t+1} . If both values were dependent on the embedded human capital h_{t+1} , the term $(1 - \rho)\mathcal{B} \ln(g(\lambda_t, \varepsilon_{t+1}))$ would disappear. The value of quitting would still depend on ρ but it would not affect the optimal λ_t .

The constraint that insures that the manager chooses the optimal investment is

$$\begin{aligned} \alpha \ln(1 - \lambda_t) + \beta E_t Q_{t+1}\left(g(\lambda_t, \varepsilon_{t+1})h_t\right) \geq \\ \alpha \ln(1 - \hat{\lambda}_t) + \beta E_t \left[(1 - \rho) \cdot \underline{Q}_{t+1}(h_t) + \rho \cdot \overline{Q}_{t+1}\left(g(\hat{\lambda}_t, \varepsilon_{t+1})h_t\right) \right], \end{aligned}$$

where λ_t is the investment recommended by the optimal contract and $\hat{\lambda}_t$ is the investment chosen by the manager (the deviation). Normalizing, we can rewrite the incentive-compatibility constraint as

$$\begin{aligned} \alpha \ln(1 - \lambda_t) + \beta E_t \left[q_{t+1} + \mathcal{B} \ln\left(g(\lambda_t, \varepsilon_{t+1})\right) \right] \geq \\ \alpha \ln(1 - \hat{\lambda}_t) + \beta E_t \left[(1 - \rho)\underline{q} + \rho\bar{q} + \rho\mathcal{B} \ln\left(g(\hat{\lambda}_t, \varepsilon_{t+1})\right) \right]. \quad (16) \end{aligned}$$

We can now provide a more explicit characterization of the manager's optimal investment deviation $\hat{\lambda}_t$. This maximizes the expected value of quitting net of the effort cost, that is, the right-hand-side of (16). Using $g(\lambda, \varepsilon) = 1 + \lambda\varepsilon$ and remembering that $\varepsilon \in \{0, 1\}$ with probabilities $1 - p$ and p , the optimality condition (6) can be written as

$$\frac{\alpha}{1 - \hat{\lambda}_t} \geq \frac{\rho\beta\mathcal{B}p}{1 + \hat{\lambda}_t}, \quad (17)$$

which is satisfied with equality if $\hat{\lambda}_t > 0$. As implied by Proposition 2, we can now see more explicitly that $\hat{\lambda}$ is increasing in the probability ρ . Therefore, when the manager faces better outside options, the strategic incentive to innovate increases.

The original contractual problem (8) with one-sided limited commitment can be reformulated in normalized form using the 'promised utility' approach. This maximizes

the normalized investor's value subject to the normalized promise-keeping, limited enforcement and incentive-compatibility constraints, that is,

$$v(q) = \max_{\lambda, c, q(\varepsilon')} \left\{ \beta y(\lambda) - c + \beta E g(\lambda, \varepsilon') v(q(\varepsilon')) \right\} \quad (18)$$

subject to (14), (15), (16).

The solution provides the investment policy $\lambda = \varphi^\lambda(q)$, the consumption policy $c = \varphi^c(q)$, and the continuation utilities $q(\varepsilon) = \varphi^q(q, \varepsilon)$. Because of the normalization, these policies are independent of h . However, once we know h , we can reconstruct the original, non-normalized values, that is, $C = ch$ and $Q = q + \mathcal{B} \ln(h)$. Also, once we know the investment policy λ and the realization of the shock ε' , we can determine the next period human capital as $(1 + \lambda \varepsilon')h$ and construct the whole sequence of human capital. Therefore, to characterize the optimal contract we can focus on the normalized policies.

The optimal policies $\varphi^\lambda(q)$, $\varphi^c(q)$, and $\varphi^q(q, \varepsilon)$ satisfy the first order conditions

$$c = \mu, \quad (19)$$

$$\frac{\alpha(\mu + \chi)}{1 - \lambda} - \beta y_\lambda(\lambda) = \beta p \left[v(q(\bar{\varepsilon})) + \frac{\mathcal{B}[\mu + \chi + (1 - \rho)\gamma(\bar{\varepsilon})]}{1 + \lambda} \right], \quad (20)$$

$$\mu(\varepsilon) = \frac{\mu + \chi + \gamma(\varepsilon)}{1 + \lambda \varepsilon}. \quad (21)$$

These conditions are derived in Appendix C. The variables μ , $\gamma(\varepsilon)$ and χ are the Lagrange multipliers for constraints (14)-(16). The envelope condition $v'(q) = -\mu$ (also derived in Appendix C) shows the equivalence between the normalized problem (18) and the original problem (8).

For the case with double-sided limited commitment, we can reformulate problem (12) in normalized form in a similar fashion. Using the 'promised utility' approach, the partnership contract with double-sided limited commitment can be written as

$$v(q) = \max_{c, q(\varepsilon)} \left\{ \beta y(\hat{\lambda}) - c + \beta E g(\hat{\lambda}, \varepsilon) v(q(\varepsilon)) \right\} \quad (22)$$

subject to

$$q = \ln(c) + \alpha \ln(1 - \hat{\lambda}) + \beta E \left[\mathcal{B} \ln(g(\hat{\lambda}, \varepsilon)) + q(\varepsilon) \right]$$

$$q(\varepsilon) = (1 - \rho) \underline{q} + \rho \bar{q} - (1 - \rho) \mathcal{B} \ln(g(\hat{\lambda}, \varepsilon)), \quad \text{for all } \varepsilon,$$

where $\hat{\lambda}$ is determined by condition (17).

In this case, the optimal deviation $\hat{\lambda}$ is independent of q . As a result, $\hat{\lambda}$ determines a lower bound on the normalized utility, denoted by q_{min} , which satisfies the condition

$$q_{min} = \ln \left(c(q_{min}) \right) + \alpha \ln(1 - \hat{\lambda}) + \beta E_t \left[(1 - \rho) \underline{q} + \rho \bar{q} + \rho \mathcal{B} \ln \left(g(\hat{\lambda}, \varepsilon) \right) \right]. \quad (23)$$

Problem (22) can be seen as a special case of problem (18) where we have replaced the incentive-compatibility constraint (16) with $\lambda = \hat{\lambda}$. Furthermore, we have imposed that the enforcement constraint (15) is always satisfied with equality. This is because any promise that exceeds the outside value of the manager will be renegotiated ex-post. Notice that in this problem the decision variables, c and $q(\varepsilon)$, are fully determined by the promise-keeping and incentive-compatibility constraints. Therefore, the problem can be solved without performing any optimization, besides solving for $\hat{\lambda}$.

2.4 Contract properties

In this subsection we illustrate the properties of the optimal contract numerically. The specific parameter values will be described in Section 4.1 where we conduct a quantitative analysis with the general model. The computational procedure used to solve the optimal contract is described in Appendix E¹¹.

As we have seen, the solution to the contractual problem (18) with one-sided limited commitment provides the optimal policies for investment, $\lambda = \varphi^\lambda(q)$, manager's consumption, $c = \varphi^c(q)$, and continuation utilities, $q(\varepsilon) = \varphi^q(q, \varepsilon)$. Because of the normalization, these policies are independent of h . However, once we know the normalized policies and the initial human capital h_0 , we can construct the whole sequence of h as well as the non-normalized values of consumption, $C = ch$, and lifetime utility, $Q = q + \mathcal{B} \ln(h)$. Therefore, to characterize the optimal contract we can focus on the normalized policies as characterized by the first order conditions (19)-(21). This is also the case for the solution to problem (22) in the environment with double-sided limited commitment.

The dynamics of promised utilities. The top panels of Figure 4 plot the values of next period normalized continuation utilities, $q(\varepsilon) = \varphi^q(q, \varepsilon)$, as functions of current normalized utility, q , for the environments with one-sided and double-sided limited commitment. We have also plotted the 45 degree line which allows us to see more clearly the dynamics of the contract in response to the shock (if the continuation utility is below (above) the 45 degree line then the next period q is smaller (bigger) than the current q). Finally, the vertical lines indicate the initial normalized values of the contract for the manager, q_0 . At this stage we have not specified yet as the initial values are determined in the two environments. These will be described in later when we embed the model in a general equilibrium. For the moment we take them as exogenous.

We discuss first the case with one-sided limited commitment. The contract starts with an initial \bar{q} indicated by the vertical line. Then, if the investment does not succeed ($\varepsilon = 0$), the next period value of q remains the same. If the investment succeeds ($\varepsilon = 1$),

¹¹Without loss of generality, we assume for the rest of the paper that $\bar{\varepsilon} = 1$.

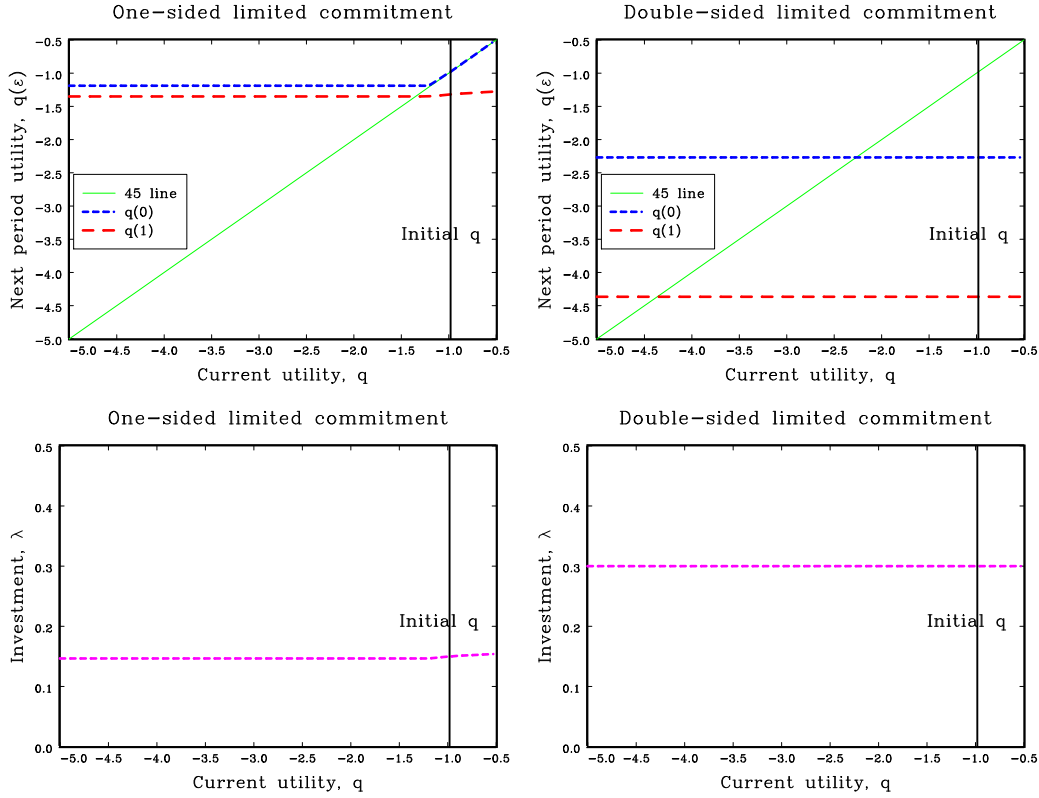


Figure 4: Continuation utilities and investment with one-sided and double-sided limited commitment.

the next period q declines until it reaches a lower bound. It is important to remember, however, that these are normalized utilities. Therefore, the fact that q declines does not necessarily mean that the actual (non-normalized) utility $Q = q + \mathcal{B} \ln(h)$ declines. For example, it could be possible that Q increases but less than $\mathcal{B} \ln(h)$.

The dynamics of promised utilities can be explained as follows. For relatively high values of q , the limited commitment constraint is not binding and the manager's value evolves as if the contract was fully enforceable. In this case it becomes optimal to provide full insurance to the manager, that is, to keep the non-normalized utility Q constant. In terms of normalized utility this means that $q = Q - \mathcal{B} \ln(h)$ remains constant when the investment fails ($\varepsilon = 0$) since in this case h does not change. When the investment succeeds ($\varepsilon = 1$), however, h increases. Then $q = Q - \mathcal{B} \ln(h)$ must fall in order to keep the non-normalized utility Q constant. However, as q declines, the enforcement constraint becomes binding. In fact, a declining q means that the non-normalized utility Q stays constant but the outside value increases with h . Eventually, the normalized utility reaches a lower bound which is indicated by the intersection of the dashed line $q(1)$ with the 45 degree line. After that the continuation utilities oscillate between two points corresponding to the intersections of the two dashed lines with the 45 degree lines.

To summarize, the contract starts with an initial normalized utility q_0 indicated by the vertical line. Then, if the realization of the shock is low, q does not change. If the realization of the shock is high, q declines until it reaches a lower bound. At this point the normalized continuation utility fluctuates between two values indicated in the graph by the intersection of the dashed lines with the 45 degree line.

The optimal policy in the environment with double-sided limited commitment is shown in the second panel of Figure 4. In this environment the investor does not commit to the contract and renegotiates any promises that exceed the outside value of the manager. As a result, the manager always receives the outside value. The only exception is in the first period when the manager receives the value indicated in the figure by the vertical line. After the initial period, q jumps immediately to the outside option and fluctuates between two values. The fact that the initial q (indicated by the vertical line) is bigger than future values implies that in the first period the manager receives a higher payment (consumption) relative to her human capital.

Investment. The bottom panels of Figure 4 plot the investment policy λ . In the environment with one-sided limited commitment, the enforcement constraint is not binding for high values of q . As a result, λ is only determined by the investment cost, part of which is given by the effort dis-utility. For lower values of q , however, the enforcement constraint for the manager is either binding or close to be binding. Consequently, a higher value of λ increases the outside value for the manager and must be associated with a higher promised utility. Since this is costly for the investor, the optimal λ is lower for low values of q (although quantitatively the dependence is small).

In the environment with double-sided limited commitment λ is independent of q since the manager always chooses $\lambda = \hat{\lambda}$. Given the limited commitment of the investor, the manager knows that the value of the contract will always be reneged to her outside value. Thus, the objective of the manager is to choose the investment that maximizes the outside value net of the utility cost of effort. But in doing so, the manager does not take into account that investment also reduces production.

For the particular parametrization considered here, the investment chosen with double-sided limited commitment is greater than in the environment with one-sided limited commitment. However, this property is not general because there are two contrasting effects. On the one-hand, with double-sided limited commitment, the manager does not take into account the loss of production when choosing the investment that maximizes the outside option. This leads to a higher choice of λ . On the other, the outside option is the value of finding employment in another firm, which happens with probability $\rho < 1$. Instead, when λ is chosen to maximize the surplus of the existing contract—which is the case in the one-sided limited commitment—the innovation adds value with probability 1. This leads to a lower choice of λ . Therefore, to have that the investment in the double-sided limited commitment is bigger than the investment with one-sided limited commitment, we need that the marginal production loss from innovation (the derivative of $y(\lambda)$) and the probability of finding another occupation (the probability ρ) are sufficiently large.

3 General model

We now embed the financial sector in a general equilibrium framework. This allows us to endogenize the parameter ρ and the outside values $\underline{Q}_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$.

There are two sectors in the model—financial and nonfinancial—and two types of agents—a unit mass of *investors*, a unit mass of *workers*. Innovations as described earlier take place only in the financial sector. This does not mean that there are no innovations outside of the financial sector. Instead, we should interpret them as ‘differential’ innovations compared to the rest of economy which, for simplicity, we do not model explicitly. An alternative interpretation of the model is that the financial sector encompasses all the ‘innovative segments’ of the economy, financial and nonfinancial, where similar organizational changes have taken place. In this paper we prefer to focus on the financial sector because the organizational and economic changes described in the introduction have been more evident.

Investors are the owners of firms and are risk neutral. The risk neutrality can be rationalized by the ability of investors to diversify their ownership of firms. Workers have the same utility $\ln(c_t) + \alpha \ln(1 - \lambda_t)$. We assume that only managerial occupations in the financial sector require effort λ_t and, therefore, the utility of workers employed in the nonfinancial sector reduces to $\ln(c_t)$. We interpret λ_t as the differential innovation effort compared to the rest of the economy.

All agents discount future utility by the factor $\hat{\beta}$ and survive with probability $1 - \omega$. In every period there are newborn agents of each type so that the population size and composition remain constant over time. Newborn workers are endowed with initial human capital h_0 . The motivation for adding this particular demographic structure is to prevent the distribution of h_t from becoming degenerate. The assumption of a constant initial human capital h_0 together with the finite lives of workers guarantee that the distribution of h_t across financial managers converges to an invariant distribution and the model is stationary in levels.

Taking into account the survival probability, the ‘effective’ discount factor is $\beta = \hat{\beta}(1 - \omega)$. Using the effective discount factor β , the previous characterization of the optimal contract between managers and investors applies to the general model without modification.

A fraction ψ of workers are born with the ability or skills to become managers in the financial sector. We denote by S the total mass of workers employed in the nonfinancial sector (with and without the ability to become financial managers) and $1 - S$ is the mass of workers employed in the financial sector. The assumption that only a fraction ψ of workers have the ability to become financial managers is only important for the quantitative properties of the model, it does not affect its qualitative properties.

The nonfinancial sector is competitive and produces output with the technology $F(H) = zH$, where z is a constant and H is the aggregate efficiency-units of labor supplied by workers employed in the nonfinancial sector. This results from the aggregation of human capital of all workers employed in the nonfinancial sector. As we will see, in equilibrium, the human capital of each worker employed in the nonfinancial sector is h_0 . Therefore, $H = h_0S$. For simplicity, we abstract from capital accumulation.

Because of the competitiveness, the wage rate (per unit of human capital) earned in the nonfinancial sector is equal to the marginal productivity, which is equal to z .

While the nonfinancial sector is competitive, the hiring process in the financial sector is characterized by matching frictions. Workers with the ability to become financial managers, find occupation in the financial sector if matched with vacancies funded by investors. Denote by ρ_{t+1} the matching probability. Then the lifetime utility of a worker currently employed in the nonfinancial sector with human capital h and with the ability to become a financial manager is

$$\underline{Q}_t(h) = \ln(h) + \beta \left[(1 - \rho_{t+1}) \cdot \underline{Q}_{t+1}(h) + \rho_{t+1} \cdot \overline{Q}_{t+1}(h) \right]. \quad (24)$$

The worker consumes the wage income h in the current period. In the next period, with probability ρ_{t+1} she finds a job in the financial sector. In this case the lifetime utility is $\overline{Q}_{t+1}(h)$. With probability $1 - \rho_{t+1}$ she remains employed in the nonfinancial sector and the lifetime utility is $\underline{Q}_{t+1}(h)$. In this extended model, the value for a skilled worker (manager) of not finding an occupation in the financial sector is the value of being employed in the nonfinancial sector. The function $\overline{Q}_{t+1}(h)$ is the value of a new contract for the worker.

3.1 Matching and general equilibrium

In the financial sector, investors post vacancies that specify the level of human capital h and the value of the contract for the manager $\overline{Q}_t(h)$. This is the value of the long-term contract signed between the firm and the manager. The cost of posting a vacancy is τh .

Let $X_t(h, \overline{Q}_t)$ be the number of vacancies posted for managers with human capital h that offer $\overline{Q}_t(h)$. Furthermore, denote by $U_t(h, \overline{Q}_t)$ the number of workers with human capital h in search of an occupation in the financial sector with posted value $\overline{Q}_t(h)$. The number of matches is determined by the matching function $m_t(h, \overline{Q}_t) = AX_t(h, \overline{Q}_t)^\eta U_t(h, \overline{Q}_t)^{1-\eta}$. The probabilities that a vacancy is filled and a worker finds occupation are $\phi_t(h, \overline{Q}_t) = m_t(h, \overline{Q}_t)/X_t(h, \overline{Q}_t)$ and $\rho_t(h, \overline{Q}_t) = m_t(h, \overline{Q}_t)/U_t(h, \overline{Q}_t)$.

Investors can freely post vacancies, giving rise in equilibrium to the free-entry condition $\phi_t(h, \overline{Q}_t) V_t(h, \overline{Q}_t) = \tau h$. The free entry condition must be satisfied for any level of human capital h .

We can now take advantage of the properties of the optimal contract characterized in the previous sections where we have shown that the value of the contract for the investor is linear in h , that is, $V_t(h, \overline{Q}_t) = v_t(\bar{q}_t)h$. The variable \bar{q}_t is the normalized value of the contract for a newly hired worker. To determine \bar{q}_t we need only to define a menu of posted contracts for all possible levels of human capital h . More precisely, once \bar{q}_t is decided, the investor offers $\overline{Q}_t = \bar{q}_t + \mathcal{B} \ln(h)$ to the worker with human capital h . Then, focusing on a symmetric equilibrium in which the probability of filling a vacancy is independent of h , the free-entry condition can be rewritten in normalized form as

$$\phi_t(\bar{q}_t) v_t(\bar{q}_t) = \tau. \quad (25)$$

Appendix D discusses the equilibrium conditions in more detail and shows that the worker receives a fraction $1 - \eta$ of the matching surplus. This is the standard efficiency

property of models with directed search. As is well known, the same outcome would arise if we assume Nash bargaining with the bargaining power of managers equal to $1 - \eta$ (the Hosios (1990) condition).

Next we normalize the employment value of workers employed in the nonfinancial sector, equation (24). This can be rewritten as

$$\underline{q}_t = \ln(1) + \beta \left[(1 - \rho_{t+1}) \cdot \underline{q}_{t+1} + \rho_{t+1} \cdot \bar{q}_{t+1} \right]. \quad (26)$$

The values \underline{q}_t and \bar{q}_t correspond to the normalized outside values used in the previous characterization of the optimal contract. The only difference is that in a general equilibrium these values could be time dependent. We now have all the ingredients to define a *steady state* general equilibrium where these values are constant.

Definition 2 (Steady state) *Given a contractual regime (one-sided or double-sided limited commitment), a stationary equilibrium is defined by*

1. Policies $\lambda = \varphi^\lambda(q)$, $c = \varphi^c(q)$, $q(\varepsilon) = \varphi^q(q, \varepsilon)$ for contracts in the financial sector;
2. Normalized utilities for workers employed in the nonfinancial sector, \underline{q} , workers newly hired in the financial sector, \bar{q} , and initial normalized value for investors, \bar{v} ;
3. Number of workers in the nonfinancial sector, S , of which U have managerial skills. Posted vacancies, X , filling probability, ϕ , and finding probability, ρ ;
4. Distribution of workers employed in the financial sector $\mathcal{M}(h, q)$;
5. Law of motion for the distribution of financial workers, $\mathcal{M}_{t+1} = \Phi(\mathcal{M}_t)$;

Such that

1. The policy rules $\varphi^\lambda(q)$, $\varphi^c(q)$, $\varphi^q(q, \varepsilon)$ solve the optimal contract;
2. The normalized utilities \underline{q} and \bar{q} and investor value \bar{v} solve (25), (26) and (37);
3. Filling and finding probabilities satisfy $\phi = m(X, U)/X$ and $\rho = m(X, U)/U$.
4. The law of motion $\Phi(\mathcal{M})$ is consistent with contract policies $\varphi^\lambda(q)$ and $\varphi^q(q, \varepsilon)$.
5. The distribution of managers is constant, that is, $\mathcal{M} = \Phi(\mathcal{M})$.

For the later analysis, it will be convenient to state formally the property for which increasing competition for managers redistributes rents in their favor. The proof is provided in Appendix D.

Lemma 3 *An increase in ρ results in a higher steady-state contract value \bar{q} offered to the manager; i.e. $\bar{q}'(\rho) > 0$.*

3.2 Inequality

The general model features two types of occupations: workers employed in the nonfinancial sector (some of whom have the ability to become managers of financial firms) and skilled workers employed in the financial sector. This permits us to study the inequality of the incomes earned across the two sectors and the inequality within the financial sector. Here we focus on the distribution within the financial sector.

Since the income of workers employed in the financial sector is proportional to human capital, we can use h as a proxy for the distribution of income. As a specific index of inequality we use the square of the coefficient of variation in human capital, that is,

$$\text{Inequality index} \equiv \frac{\text{Var}(h)}{\text{Ave}(h)^2}.$$

In a steady state equilibrium with double-sided limited commitment, the inequality index can be calculated exactly. Let's first derive the steady state employment in the financial sector, $1 - S$. This can be derived from the flow of workers with managerial ability into financial occupations (at rate ρ) and out of financial occupations (at rate ω), that is, $1 - S_{t+1} = (1 - S_t)(1 - \omega) + U(1 - \omega)\rho_{t+1}$. The equivalent equation for workers with managerial ability is $U_{t+1} = U_t(1 - \omega)(1 - \rho) + \omega\psi$. After imposing steady state conditions, these two equations can be solved for the stock of workers employed in the financial sector,

$$1 - S = \frac{\rho(1 - \omega)\psi}{\rho + \omega - \rho\omega}.$$

Next we compute the average human capital for the mass $1 - S$ of workers employed in the financial sector,

$$\text{Ave}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j h_j.$$

The index j denotes the employment tenure for active managers (employment periods). Therefore, $j = 0$ identifies newly hired workers. Since managers survive with probability $1 - \omega$, the fraction of managers who have been active for j periods is $\omega(1 - \omega)^j$.

The variance of h across the $1 - S$ workers is calculated as

$$\text{Var}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j \left(h_j - \text{Ave}(h) \right)^2,$$

which has a similar interpretation as the formula used to compute the average h .

Using the property of the model with double-sided limited commitment where all firms choose the same λ and, therefore, all managers experience the same expected growth in human capital, Appendix F shows that the average human capital and the inequality index take the forms

$$\text{Ave}(h) = h_0 \left[\frac{\omega}{1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)} \right], \quad (27)$$

$$\text{Inequality index} = \frac{[1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)]^2}{\omega[1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)]} - 1. \quad (28)$$

The average human capital and the inequality index are simple functions of the investment $\hat{\lambda}$. We then have the following proposition.

Lemma 4 *The average human capital and the inequality index for financial managers are strictly increasing in $\hat{\lambda}$.*

That average human capital increases with investment is obvious. The dependence of inequality on $\hat{\lambda}$ can be explained as follows. If $\hat{\lambda} = 0$, the human capital of all managers will be equal to h_0 and the inequality index is zero. As $\hat{\lambda}$ becomes positive, inequality increases for two reasons. First, since the growth rate $g(\hat{\lambda}, \varepsilon)$ is stochastic, human capital will differ *within* the same tenure cohort of managers (managers with the same employment tenure). Second, since each cohort experiences growth, the average human capital differs *between* cohorts of managers. More importantly, the cross sectional dispersion in human capital induced by these two mechanisms (the numerator of the inequality index) dominates the increase in average human capital (the denominator of the inequality index). Thus, inequality increases in $\hat{\lambda}$.

We can compute explicitly the *within* and *between* cohort inequality by decomposing the variance of h as follows:

$$\text{Var}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j \left(h_j - \text{Ave}_j(h) \right)^2 + \omega \sum_{j=0}^{\infty} (1 - \omega)^j \left(\bar{h}_j - \text{Ave}(h) \right)^2,$$

where $\text{Ave}_j(h)$ is the average human capital for the j cohort (managers employed for j periods). The first term sums the variances of each cohort while the second term sums the squared deviation of each cohort from the overall average. Using the above decomposition, the appendix shows that the *within* and *between* cohort inequality indices have simple analytical expressions and they are both strictly increasing in $\hat{\lambda}$.

4 The impact of organizational changes

We now explore the core issue addressed in this paper, that is, how the organizational change described in the introduction affects risk taking, sectoral income, valuation of financial firms and inequality. We have identified two key effects from the organizational change in the financial sector:

1. ***Increased competition for managers:*** The separation between investors and managers expanded the base of potential investors who could fund new investment projects, facilitating the creation of new businesses. In the context of our model this is captured, parsimoniously, by a reduction in the vacancy cost τ . A lower value of τ generates the creation of more vacancies and, therefore, more competition for managers.
2. ***Weakened the commitment of investors:*** While the limited commitment of managers was also a feature of the traditional partnership (managers were not

prevented from leaving the partnership), the commitment of investors was much stronger since there was not a sharp distinction between investors and managers. Even from a legal stand point, it was difficult for a partnership to replace a partner without a consensual agreement. A feature of a corporation, instead, is a clearer separation between investors and managers. In the context of our model, this is captured by a shift from the environment with one-sided limited commitment to the environment with double-sided limited commitment.

In summary, we formalize the demise of the traditional partnership as a shift to an environment where there is more competition for managers (it is easier to fund new business managed by financial managers) and where contracts have limited enforceability for both investors and managers. We explore first the consequences of greater competition for managers in the environment with double-sided limited commitment.

Proposition 5 *In the environment with double-sided limited commitment, a steady state equilibrium with a lower value of τ features:*

1. *Greater risk-taking, that is, higher $\hat{\lambda}$.*
2. *Bigger size and higher relative productivity of the financial sector.*
3. *Lower stock market valuation of financial firms.*
4. *Greater income inequality between sectors (financial and nonfinancial) and within the financial sector.*

The first property is an immediate consequence of Proposition 2: the lower value of τ increases the probability of a match and, consequently, it raises the incentive of managers to exert more effort to increase their outside value.

The increase in the size of the financial sector derives in part from higher employment and in part from higher investment. We would like to point out that, although the increase in the share of employment would arise even if there were no contractual frictions, the increase in investment would only arise only with contractual frictions. This is a novel feature of our model which is key to capture the ‘productivity’ increase in the financial sector relatively to other sectors, consistent with the pattern shown in Figure 1. According to that figure, the share of value added of the financial sector has increased much more than the share of employment.

The third property—lower valuation of financial firms—is a direct consequence of Lemma 3: the initial value of the contract for the manager, \bar{q} , increases with the probability of a match ρ , which is higher in the steady state with a lower value of τ (as already mentioned above). This effect of increased competition for managers is common across organizational forms in which there is a division between investors and managers, even if contracts were fully enforceable. However, the effect is likely to be stronger when there is limited commitment also for investors. This will be shown numerically in the quantitative simulation.

Finally, the fourth property—greater inequality—follows from the first property, that is, from the higher investment $\hat{\lambda}$. As we have seen in Lemma 4, a higher value of $\hat{\lambda}$ increases human capital accumulation and inequality within the financial sector. At the same time, since workers that remain employed in the nonfinancial sector do not accumulate human capital while the human capital of workers employed in the financial sector grow faster on average, we have greater income inequality between the two sectors.

The next question is how the equilibrium properties are affected by the second implication of the structural change, that is, a shift from an environment with one-sided limited commitment to an environment with double-sided limited commitment. We characterize the effects numerically since the consequences of this shift cannot be characterized analytically.

4.1 Quantitative analysis

We calibrate the model annually using data for the 2000s. Since in the 2000s the partnership form of organization was no longer dominant in the financial sector, we calibrate the model under the environment with double-sided limited commitment.

The only functional form that has not been specified is the production function in the financial sector. We assume a quadratic form, that is, $y(\lambda) = 1 - \lambda^2$. Therefore, if a worker devotes all of her time to production ($\lambda = 0$), each unit of human capital produces one unit of output. If instead the worker allocates all of her time to innovating ($\lambda = 1$), production is zero.

Given the specification of preferences and technology and after normalizing the initial human capital h_0 to 1, there are 9 parameters to calibrate (see the top section of Table 1). Given the difficulty of calibrating the parameter of the matching function η , it is customary to set it to $\eta = 0.5$. We follow the same approach here even though in our model jobs are created through matching only in the financial sector. We are then left with 8 parameters which we calibrate using the 8 moments listed in the bottom section of Table 1.

The first 5 moments come from direct empirical observations or typical calibration targets. An interest rate of 4% is standard in the calibration of macroeconomic models. A lifetime of 40 years corresponds to an approximate duration of working life. The employment and value added shares are the approximate numbers for finance and insurance in the 2000s as shown in Figure 1. The inequality index comes from the 2010 Survey of Consumer Finance for the sample of managerial occupations in the financial sector (see Figure 2 for a more detailed description of the data). The last three moments (innovation time, job finding rate and job filling rate) are not based on direct empirical observations and the values assigned are somewhat arbitrary. A sensitivity analysis will clarify the relevance of these calibration targets. Appendix G provides a detailed description of how the 8 moments are mapped into the 8 parameters.

Results. Our goal is to assess the quantitative impact of greater competition and lower contract enforcement. The impact of higher competition is captured by looking at the equilibrium consequences of reducing the vacancy cost τ . The impact of lower enforce-

Table 1: Parameters and calibration moments.

<i>Parameters</i>		
$\hat{\beta}$	Discount factor	0.962
ω	Death probability	0.025
z	Productivity in the nonfinancial sector	0.731
ψ	Fraction of workers searching for financial jobs	0.042
p	Probability of successful innovation	0.035
α	Utility parameter for dis-utility innovation effort	0.139
τ	Cost of posting a vacancy in the financial sector	0.174
A	Matching productivity	0.500
η	Matching share parameter (pre-set)	0.500
<i>Calibration moments</i>		
	Interest rate	0.04
	Life expectancy of workers	40.00
	Employment share in finance	0.04
	Value added share in finance	0.08
	Inequality index (coeff. variation) in financial sector	2.00
	Time allocated to innovation in finance	0.30
	Probability of finding an occupation in finance	0.50
	Probability of filling a vacancy	0.50

ment is captured by looking at the changes induced by a shift from the environment with one-sided limited commitment to the environment with double-sided limited commitment. We see the environment with one-sided limited commitment and higher vacancy cost as characterizing the financial sector in the pre-1980s period. The environment with double-sided limited commitment and lower vacancy cost is representative of recent years.

Since the vacancy cost τ has been calibrated using the 2000s data, for the pre-1980s period we have to assign a higher number that, ideally, we would like to pin down using some calibration target. Since it is difficult to identify such a target, we start with the assumption that in the pre-1980s period the cost was 50% higher.

Figure 5 plots the steady state policy $\lambda = \varphi^\lambda(q)$ in the environments with one-sided and double-sided limited commitment, and for two values of τ . In the environment with one-sided limited commitment, more competition (lower τ) reduces slightly the investment λ (although the change is so small that it is difficult to see in the graph). This is because, as shown in Table 2, the probability of receiving offers increases with more competition. Since this raises the outside value of managers, a larger share of the return must be shared with managers, making the investment less attractive for investors. All of this is consistent with Proposition 1.

In contrast, when neither managers nor investors can commit, more competition induces more innovation, as Proposition 2 predicts. Also in this environment the probability of external offers increases, which raises the external value of managers and makes investment less attractive for investors. In order to implement the optimal λ , investors would need to promise adequate future compensation. The problem is that future promises are not credible with double-sided limited commitment and the only way managers can in-

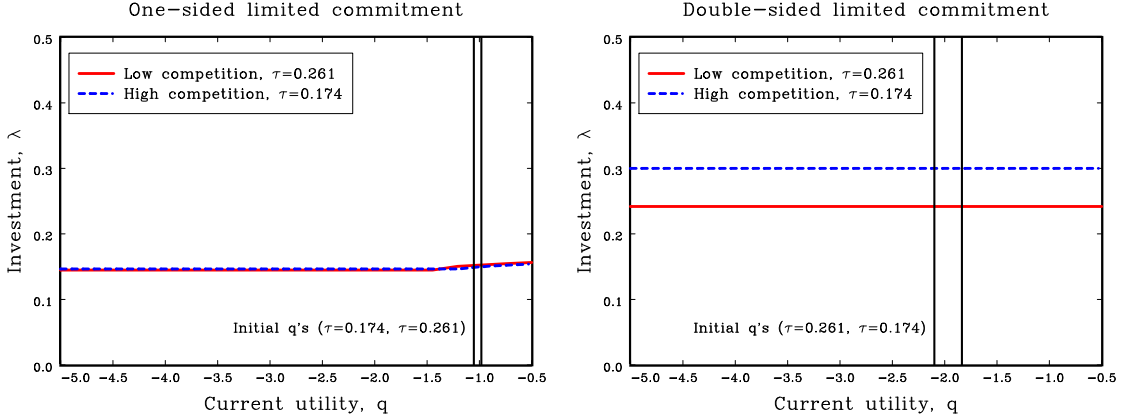


Figure 5: Steady state investment policies for different τ in the environments with one-sided and double-sided limited commitment.

crease their contract value is by raising their outside value. This is achieved by choosing higher λ . With a lower τ , the probability of an external offer ρ increases. Since the manager benefits from higher innovation only if she receives an external offer, the higher probability ρ raises the manager's incentive to choose a higher value of λ .

So far we have shown that the organizational change that took place in the financial sector induced more risk-taking. We now show that they also generated other changes that are consistent with the observations we highlighted in the introduction. Table 2 shows that the shift to an environment with double-sided limited commitment and lower τ is associated with an insignificant change in the share of employment in the financial sector but a significant increase in the share of output. The output share increases from 6.5% to 8%.

Another important prediction of the model is that the shift is associated with a reduction in the (average) value of investors, relative to human capital. Since we do not have physical capital, we use human capital as a proxy for the book value of assets.¹² Table 2 also shows that the initial investor's value and the probability of filling a vacancy are both lower. This follows directly from Lemma 3 and the free entry condition $\phi(\bar{q}) \cdot v(\bar{q}) = \tau$ after the reduction in the vacancy cost τ .

Table 2 also shows why the investor's commitment to a long-term contract can be weakened by competition. As expected, an increase in competition for managers results in a redistribution in favour of the managers, independently of the level of commitment. However, at any level of competition, a move from one-sided to two-sided limited commitment increases the normalised *ex-post* value of the investor, $Ev(q)$; and, even more, the non-normalized *ex-post* value since growth is higher. Therefore, the investor may be tempted to recover his *ex-post* relative losses due to increased competition by renegeing on his commitments. Such a move to a double-sided limited commitment economy may re-

¹²This would be the case if we explicitly introduce capital and assume that there is complementarity between human and physical capital.

Table 2: Steady state properties of equilibria associated with different values of τ in the environments with one-sided and double-sided limited commitment.

	<i>One-sided limited commitment</i>	<i>Double-sided limited commitment</i>
Low competition ($\tau = 0.261$)		
Average value of λ	0.151	0.242
Offer probability, ρ	0.445	0.441
Filling probability, ϕ	0.561	0.567
Share of employment financial sector	0.040	0.040
Share of output financial sector	0.065	0.073
Earnings in the nonfinancial sector	0.731	0.731
Earnings in the financial sector	1.110	1.257
Initial investor value \bar{v}	0.464	0.460
Average investor value $Ev(q)$	0.581	0.716
Within inequality fin sector	0.056	0.369
Between inequality fin sector	0.071	0.313
Coefficient of variation	0.356	0.826
High competition ($\tau = 0.174$)		
Average value of λ	0.147	0.300
Offer probability, ρ	0.497	0.500
Filling probability, ϕ	0.503	0.500
Share of employment financial sector	0.040	0.040
Share of output financial sector	0.065	0.080
Earnings in the nonfinancial sector	0.731	0.731
Earnings in the financial sector	1.116	1.388
Initial investor value \bar{v}	0.388	0.348
Average investor value $Ev(q)$	0.442	0.537
Within inequality fin sector	0.054	3.110
Between inequality fin sector	0.069	0.890
Coefficient of variation	0.351	2.000

duce the investor's initial value (as Table 2 shows), but definitively increase his expected value *ex-post*.

The above properties are consistent with the observed expansion of the financial sector and the decline in market valuation of financial institutions, relative to other sectors, as shown in Figure 3. The model also generates an increase in income inequality between the financial and nonfinancial sectors and within the financial sector, consistent with the evidence provided in Figures 1 and 2.

5 Conclusion

The financial crisis of 2007-2009 has brought attention to the growth in size and importance of the financial sector over the past few decades, as well as the increase in risk taking by financial managers. Much attention has also been placed on the extremely high

compensation of financial professionals. Why did these trends emerge over this period of time? In this paper we have argued that changes in the organizational structure of financial firms have increased competition for managerial skills and weakened the enforcement of contractual relationships between managers and investors¹³. These changes could have also played an important role in another widely documented trend occurred during the same period—the increase in income inequality.

The fact that inequality has increased over time, especially in anglo-saxon countries, is well documented (e.g. Saez and Piketty (2003)). The increase in inequality has been particularly steep for managerial occupations in financial industries (e.g. Bell and Van Reenen (2010)). In this paper we propose one possible explanation for this change. We emphasize the increase in competition for human talent that followed the organizational changes in the financial sector. In an industry where the enforcement of contractual relations is limited, the increase in competition raises the managerial incentives to undertake risky investments. Although risky innovations may have a positive effect on aggregate production, the equilibrium outcome may not be efficient and generates greater income inequality. The higher competition for managerial talent seems consistent with the evidence that managerial turnover, although not explicitly modelled in the paper, has also increased during the last thirty years.

We have shown these effects through a dynamic general equilibrium model with long-term contracts, subject to different levels of commitment and enforcement. The model features two sectors—financial and nonfinancial—with innovations taking place only in the financial sector. Of course, the assumption that only the financial sector innovates is a simplification that we made to keep the model tractable and the analysis focused. An alternative interpretation of the model is that the financial sector represents the collection of the most ‘innovative segments’ of the economy, financial and nonfinancial, where similar contractual frictions emerge and the type of organizational changes described in the paper could have similar effects.

In this sense, our model is general and has general prescriptions. When organizations are subject to external competition—with different effects on members of the organization—competition is likely to distort internal decisions and result in redistribution of *ex-post* rents. With enough commitment (in our model: one-sided limited commitment), the organization can internalize these distortions but this does not mean it can implement the *ex-ante* full-commitment allocation which makes the organization immune to *ex-post* competition (with one-sided limited commitment there is lower risk-taking in response to competition).

We described our framework as a model of the innovative financial sector for several reasons. First, it is in the innovative financial sector where the organizational changes described in the introduction have been more evident. Second, some of the features of this sector—that our model helps to explain—are less present in other sectors (for example, the relatively low book value). Third, as in our model, it is the financial sector where *managerial talent* is the most relevant factor of production and it is particularly *inalienable* (capital and unskilled labor play a more relevant role in other innovative

¹³See Footnotes 6 and 9 for a brief reference to alternative explanations.

sectors and patents on financial instrument are *rare avis* and difficult to enforce).¹⁴

It can be argued that modern financial organizations have many credible instruments (bonuses, etc.) to overcome the investor's commitment problem and, therefore, that our model with two-sided limited commitment is a poor description of innovative financial firms. But we have explicitly chosen to work with a simplified model in order to sharpen the key mechanism that emerges in the presence of limited commitment. Sophisticated compensation packages for CEOs and financial managers are just partial forms of limited commitment compared to the internal compensation schemes that dominated in the previous organizational form, that is, the traditional partnership.¹⁵

¹⁴Although these differences with other innovative sectors may be a question of degree “But perhaps *the most significant change has been to human capital*. Recent changes in the nature of organizations, the extent and requirements of markets, and the availability of financing have made specialized human capital much more important, and also much more mobile. But *human capital is inalienable, and power over it has to be obtained through mechanisms other than ownership*”. Rajan and Zingales (2000).

¹⁵“The highest incomes and the largest fortunes in the financial sector were made by investing one's money—in other words, as a partner of a private bank rather than as a manager of a joint stock bank.” Cassis (2013).

Appendices

A The Traditional Partnership Problem

The Traditional Partnership form does not fit easily into the contracting structure that we use to characterize incentives in modern financial firms. We model the traditional partnership as a *representative partner* of the partnership, who makes the investment decisions and who is the unique claimant of the investment rents. To make it as close as possible to our two-agent contract, we consider that the *representative partner* has the same preferences of our manager and we maintain the same timing of the decisions; however, now C_{t+1} is the consumption of representative partner at the beginning of period $t + 1$ (i.e. the partner only consumes out of her returns from the investment; except for period zero, for which we assume her consumption being given). The single agent representation implies that there are no problems of breaking the contract or distorting the investment decision; therefore, we do not need to account for incentive constraints and, to simplify, we do not model her outside options.

The *traditional partnership problem* takes the form:

$$VP(h_t) = \max_{\lambda_t} \{ \beta u(y(\lambda_t)h_t) - e(\lambda_t) + \beta E_t VP(h_{t+1}) \} \quad (29)$$

$$\text{s.t. } h_{t+1} = g(\lambda_t, \varepsilon_{t+1})h_t.$$

which results in the following Euler's equation:

$$\beta^2 E_t u'(C_{t+2}) y(\lambda_{t+1}^*) g_{\lambda}(\lambda_t^*, \varepsilon_{t+1}) h_t \leq -\beta u'(C_{t+1}) y'(\lambda_t^*) h_t + e'(\lambda_t^*), \quad (30)$$

with equality if $\lambda_t^* > 0$.

If we assume that the partner has log utility preferences for consumption, (30) simplifies to:

$$p\beta^2(1 + \lambda_t^*)^{-1} \leq -\beta \frac{y'(\lambda_t^*)}{y(\lambda_t^*)} + e'(\lambda_t^*) \quad (31)$$

If, in addition, we use the functional forms $e(\lambda) = -\alpha(1 - \lambda)$ and $y(\lambda) = 1 - \lambda^2$, then (31) simplifies to

$$p\beta^2 \leq f(\lambda) \equiv \frac{\alpha + \lambda(2 + \alpha)}{1 - \lambda}. \quad (32)$$

Since $f(0) = \alpha$ and $f'(\lambda) > 0$, it follows that if $p\beta^2 < \alpha$ then $\lambda_t^* = 0$, while if $p\beta^2 > \alpha$ then there is a unique $\lambda_t^* \in (0, 1)$. In particular, if we use the parameterisation of our calibration $p\beta^2 < \alpha$ and, therefore, the optimal investment decision for the corresponding traditional partnership is $\lambda_t^* = 0$.

B Proof of Proposition 1

In order to prove Proposition 1, first notice that the contractual Problem (9) takes the following form when it is normalised by h :

$$\begin{aligned} \min_{\chi, \gamma(\varepsilon')} \max_{c, \lambda} \quad & \left\{ \beta y(\lambda) - c + \mu \left(u(ch) - e(\lambda) \right) - \chi \left(e(\lambda) - e(\hat{\lambda}) \right) \right. \\ & + \beta E \left[v(\mu') g(\lambda, \varepsilon') + (\mu + \chi + \gamma(\varepsilon')) Q(h', \mu') \right. \\ & \left. \left. - \chi D \left(h, g(\hat{\lambda}, \varepsilon') h, \rho \right) - \gamma(\varepsilon') D(h, h', \rho) \right] \right\} \\ \text{s.t.} \quad & h' = g(\lambda, \varepsilon') h, \quad \mu' = (\mu + \chi + \gamma(\varepsilon')) / g(\lambda, \varepsilon'), \end{aligned}$$

and the corresponding first-order condition with respect to λ is given by (11):

$$\begin{aligned} (\mu + \chi) e_\lambda(\lambda) - \beta y_\lambda(\lambda) \geq \beta E \left[\left(v(\mu') + (\mu + \chi + \gamma(\varepsilon')) Q_h(h', \mu') \right. \right. \\ \left. \left. - \gamma(\varepsilon') D_2(h, h', \rho) \right) \varepsilon' \right]. \end{aligned}$$

An increase in ρ , before λ is chosen, has a direct effect on the *enforcement constraint* when $\gamma_t(\varepsilon_{t+1}) > 0$ and it is given by $D_{2,3}(h, h', \rho)$. By the definition of D , (3), $D_{2,3} > 0$ and, therefore, this direct effect of the enforcement constraint makes investment more costly. Furthermore, an increase in ρ , by making the incentive and enforcement constraints tighter, increases the value of the respective multipliers – possibly, from zero to a positive value – since $D_3 > 0$, which in turn increases μ' . The simple effect on the multipliers it's already accounted for, by the same constraints. That is, increasing χ results in $\beta E \left[Q_h(h', \mu') \varepsilon' \right] - e_\lambda(\lambda) - \beta y_\lambda(\lambda) \leq 0$, where the inequality follows from the fact that otherwise $\chi = 0$; similarly, increasing $\gamma(\varepsilon')$ results in $Q_h(h', \mu') - D_2(h, h', \rho) \leq 0$. There only remain the effects of increasing μ' , which are given by $v'(\mu') < 0$ and $Q_{h,\mu}$. Therefore if, as we assume, $Q_{h,\mu} \leq 0$, the effect of an increase in ρ is, unambiguously, a lower optimal λ^* .

Comment to Proposition 1. The assumption $Q_{h,\mu} \leq 0$ may not hold and the result of Proposition 1 remain the same, since the effect on $Q_{h,\mu}$ is likely to be dominated by the other unambiguous effects. Nevertheless, the assumption is fairly general: it only says that the increase in the manager's value due to an increase in h is not complemented by an additional increase when μ also raises. In particular, if the manager has CRRA preferences for consumption, of the form

$$u(C) = \frac{C^{1-\sigma}}{1-\sigma}$$

the optimal consumption policy, (10), takes the form: $(ch)^{-\sigma} = (h\mu)^{-1}$; that is,

$$u(h\mu) = \frac{(h\mu)^{\frac{1-\sigma}{\sigma}}}{1-\sigma},$$

and, therefore,

$$u_{h,\mu}(h\mu) = \frac{1-\sigma}{\sigma^2} (h\mu)^{\frac{1-2\sigma}{\sigma}}.$$

In sum, $u_{h,\mu}(h\mu) \leq 0$ if and only if $\sigma \geq 1$; i.e. if and only if *the intertemporal elasticity of substitution* is less or equal one. Otherwise, if $1/\sigma > 1$ the optimal contract will tend to lower current consumption in exchange for compensating the manager in the future with $Q_{h,\mu}(h\mu) > 0$. Notice that, given the separability between consumption and effort Q inherits its differentiability properties from u (we abstract from some technicalities in making this claim). We analyse in detail the particular case of $\sigma = 1$; i.e. $Q_{h,\mu}(h\mu) = 0$.

C The first-order conditions of Problem (18)

Let μ and $\gamma(\varepsilon)$ be the lagrange multipliers associated with the promise-keeping constraint and the enforcement constraint. Then the lagrangian can be written as

$$\begin{aligned} v(q) &= \beta y(\lambda) - c + \beta \sum_{\varepsilon} g(\lambda, \varepsilon) v(q(\varepsilon)) p(\varepsilon) \\ &+ \mu \left\{ \ln(c) + \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left[\mathcal{B} \ln(g(\lambda, \varepsilon)) + q(\varepsilon) \right] p(\varepsilon) - q \right\} \\ &+ \chi \left\{ \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left[\mathcal{B} \ln(g(\lambda, \varepsilon)) + q(\varepsilon) \right] p(\varepsilon) - D \right\} \\ &+ \beta \sum_{\varepsilon} \left[q(\varepsilon) - d + (1-\rho) \mathcal{B} \ln(g(\lambda, \varepsilon)) \right] \gamma(\varepsilon) p(\varepsilon) \end{aligned}$$

The first order conditions with respect to λ , c and $q(\varepsilon)$ are, respectively,

$$\beta y_{\lambda}(\lambda) + \beta \sum_{\varepsilon} \left[g_{\lambda}(\lambda, \varepsilon) v(q(\varepsilon)) + \mathcal{B} \left(\frac{g_{\lambda}(\lambda, \varepsilon)}{g(\lambda, \varepsilon)} \right) (\mu + \chi + (1-\rho)\gamma(\varepsilon)) \right] p(\varepsilon) - \frac{\alpha(\mu + \chi)}{1-\lambda} = 0$$

$$-1 + \frac{\mu}{c} = 0$$

$$g(\lambda, \varepsilon) v_q(q(\varepsilon)) + (\mu + \chi + \gamma(\varepsilon)) = 0$$

Substituting the envelope condition $v_q(q) = -\mu$ and using the functional forms of $y(\lambda)$ and $g(\lambda, \varepsilon)$ we obtain equations (20)-(21).

D The posted contract

As it is well known, with directed search there is an indeterminacy of rational expectations equilibria based on agents coordinating on arbitrary beliefs. Following the literature on directed search, we restrict beliefs by assuming that searching managers believe that small variations in matching value are compensated by small variations in matching probabilities so that the expected application value remains constant. See Shi (2006). More specifically, if $\bar{Q}_t^*(h)$ is the value of the equilibrium contract, then for any $\bar{Q}_t(h)$ in a neighbourhood $\mathcal{N}(\bar{Q}_t^*)$ of $\bar{Q}_t^*(h)$, the following condition is satisfied,

$$\rho_t(h, \bar{Q}_t(h)) \cdot [\bar{Q}_t(h) - \underline{Q}_t(h)] = \rho_t(h, \bar{Q}_t^*(h)) \cdot [\bar{Q}_t^*(h) - \underline{Q}_t(h)], \quad (33)$$

where we have made explicit that the probability of a match depends on the value received by the manager. This condition says that managers are indifferent in applying to different employers who offer similar contracts since lower values are associated with higher probabilities of matching. In a competitive equilibrium with directed search, investors take $\bar{Q}_t^*(h)$ as given and choose the contract by solving the problem

$$\max_{\bar{Q}_t(h)} \left\{ \phi_t(h, \bar{Q}_t(h)) \cdot V(h, \bar{Q}_t(h)) \right\} \quad (34)$$

subject to (33),

where $V_t(h, Q)$ is the value for the investor. The analysis of the optimal contract after matching have shown that the investor's value is a function of the value promised to the manager. The equilibrium solution also provides the initial value of the contract for the investor¹⁶ $V_t(h, \bar{Q}_t(h))$.

For any h , if $\bar{Q}_t'(h)$ is also the value of an equilibrium contract, the investor must be indifferent: $\phi_t(h, \bar{Q}_t'(h)) \cdot V_t(h, \bar{Q}_t'(h)) = \phi_t(h, \bar{Q}_t^*(h)) \cdot V_t(h, \bar{Q}_t^*(h))$. Therefore, we will only consider symmetric equilibria where investors offer the same contract (h, \bar{Q}_t) .

Furthermore, competition in posting vacancies implies that, for any level of human capital h , the following free entry condition must be satisfied in equilibrium,

$$\phi_t(h, \bar{Q}_t(h)) \cdot V_t(h, \bar{Q}_t(h)) = \tau h. \quad (35)$$

We can take advantage of the of the linear property of the model and normalize the above equations. We have shown that the value of a contract for the investor is linear in h , that is, $V_t(h, Q_t(h)) = v_t(q_t)h_t$. Therefore, the free entry condition can be rewritten in normalized form as

$$\phi_t(\bar{q}_t) \cdot v_t(\bar{q}_t) = \tau. \quad (36)$$

¹⁶Given the free entry condition, the 'initial value' for the investor is 0 and the initial value of the contract is, in fact, his 'interim value', but when there is no confusion we also refer to the initial value of the contract as the 'initial value'.

This takes also into account that we focus on a symmetric equilibrium in which the probability of filling a vacancy is independent of h (which justifies the omission of h as an explicit argument in the probability ϕ_t)¹⁷.

The investor's problem (34) can be rewritten as

$$\bar{q}_t = \arg \max_q \left\{ \phi_t(q) \cdot v_t(q) \right\}$$

subject to

$$\rho_t(q)(q - \underline{q}_t) = \rho_t(\bar{q}_t^*)(\bar{q}_t^* - \underline{q}_t), \quad \forall q \in \mathcal{N}(\bar{q}_t^*)$$

We can solve for the normalized initial utility \bar{q}_t by deriving the first order condition which can be rearranged as

$$1 - \eta = \frac{-v_t'(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}{v_t(\bar{q}_t) - v_t'(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}. \quad (37)$$

The right-hand side is the share of the surplus (in utility terms) going to the manager. Thus, the manager receives the fraction $1 - \eta$ of the surplus created by the match.

We now turn to Lemma 3, which is a special case of a more general result we prove here. Let $v_\epsilon(\bar{q})$ denote the elasticity of the investor's value function; i.e. $v_\epsilon(\bar{q}) \equiv -\frac{v_t'(\bar{q})\bar{q}}{v_t(\bar{q})}$. Our log-linear specification implies that $v_\epsilon'(\bar{q}) > 0$.

Lemma 3A $v_\epsilon'(\bar{q}) > 0$ implies $\bar{q}'(\rho) > 0$.

The optimality condition (37) can be written as

$$\frac{1 - \eta}{\eta} = v_\epsilon(\bar{q}) \frac{\bar{q} - \underline{q}}{\bar{q}}. \quad (38)$$

In a stationary equilibrium, using (26) we obtain:

$$\begin{aligned} \bar{q} - \underline{q} &= \bar{q} - \left\{ \ln(1) + \beta \left[(1 - \rho)\underline{q} + \rho\bar{q} \right] \right\} \\ &= (1 - \beta)\bar{q} + \beta(1 - \rho)(\bar{q} - \underline{q}) \\ &= (1 - \beta(1 - \rho))^{-1} (1 - \beta)\bar{q}; \end{aligned}$$

¹⁷In equilibrium only skilled workers who have never been employed in the financial sector will be actively searching. Since they have never been employed in the financial sector, they all have human capital h_0 . For determining the probability of a match when a financial manager decides to quit, we incur the problem that the number of posted vacancies is discrete. In this case we assume that investors randomize over the posting of a vacancy that is targeted at a manager with human capital h .

therefore

$$\begin{aligned} v_\epsilon(\bar{q}) &= \frac{1-\eta}{\eta} \frac{\bar{q}}{\bar{q}-\underline{q}} \\ &= \frac{1-\eta}{\eta} \frac{(1-\beta)(1-\rho)}{(1-\beta)}. \end{aligned}$$

Taking derivatives with respect to ρ ,

$$v'_\epsilon(\bar{q})\bar{q}'(\rho) = \frac{1-\eta}{\eta} \frac{\beta}{1-\beta} > 0;$$

it follows that $\bar{q}'(\rho) > 0$ if $v'_\epsilon(\bar{q}) > 0$.

E The numerical solution

We describe first the numerical procedure used to solve Problem (18) for exogenous outside values \underline{q} and \bar{q} and for exogenous probability of offers ρ . We will then describe how these variables are determined in the steady state equilibrium.

Solving the optimal contract. The iterative procedure is based on the guesses for two functions

$$\begin{aligned} \mu &= \psi(q) \\ v &= \Psi(q). \end{aligned}$$

The first function returns the multiplier γ (derivative of investor's value) as a function of the promised utility. The second function gives us the investor value v also as a function of the promised utility.

Given the functions $\psi(q)$ and $\Psi(q)$, we can solve the system

$$\beta \left[v(q(1)) + \left(\frac{\mathcal{B}}{1+\lambda} \right) (\mu + \chi + (1-\rho)\gamma(1)) \right] p = -\beta y_\lambda(\lambda) + \frac{\alpha(\gamma + \chi)}{1-\lambda} \quad (39)$$

$$c = \gamma \quad (40)$$

$$g(\lambda, \varepsilon) \psi(q(\varepsilon)) = \mu + \chi + \gamma(\varepsilon) \quad (41)$$

$$v = \beta y(\lambda) - c + \beta \sum_{\varepsilon} g(\lambda, \varepsilon) \Psi(q(\varepsilon)) p(\varepsilon) \quad (42)$$

$$q = \ln(c) + \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left(\mathcal{B} \ln(g(\lambda, \varepsilon)) + q(\varepsilon) \right) p(\varepsilon) \quad (43)$$

$$\chi \left\{ \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left[q(\varepsilon) + \mathcal{B} \ln(g(\lambda, \varepsilon)) \right] p(\varepsilon) \right. \\ \left. - \alpha \ln(1-\hat{\lambda}) - \beta \sum_{\varepsilon} \left[(1-\rho)\underline{q} + \rho\bar{q} + \rho\mathcal{B} \ln(g(\hat{\lambda}, \varepsilon)) \right] p(\varepsilon) \right\} = 0 \quad (44)$$

$$\gamma(\varepsilon) \left[q(\varepsilon) - (1-\rho)\underline{q} - \rho\bar{q} + (1-\rho)\mathcal{B} \ln(g(\lambda, \varepsilon)) \right] = 0 \quad (45)$$

The first three equations are the first order conditions with respect to λ , c , $q(\varepsilon)$, respectively. Equation (42) defines the value for the investor and equation (43) is the promise-keeping constraint. Equations (44) and (45) formalize the Kuhn-Tucker conditions for the incentive-compatibility and enforcement constraints.

Notice that equations (44) and (45) must be satisfied for all values of ε which can take two values. Therefore, we have a system of 9 equations in 9 unknowns: λ , c , v , μ , χ , $q(\varepsilon)$, $\gamma(\varepsilon)$. Once we have solved for the unknowns we can update the functions $\psi(q)$ and $\Psi(q)$ using the solutions for v and μ .

Solving for the steady state equilibrium. The iteration starts by guessing the steady state values of \bar{q} and ρ . Given these two values, we can determine \underline{q} using equation (26). With these guesses we can solve for the optimal contract as described above. This returns the functions $\mu = \psi(q)$ and $v = \Psi(q)$ in addition to $\lambda = \varphi^\lambda(q)$ and $q(\varepsilon) = \varphi^q(q, \varepsilon)$.

Once we have these functions we determine the new values of \bar{q} and ρ using the free-entry condition (36) and the bargaining condition (37). We keep iterating until convergence, that is, the guessed values of \bar{q} and ρ are equal to the computed values (up to a small approximation error).

F Derivation of the inequality index

In each period there are different cohorts of active managers who have been employed for j periods. Because managers die with probability ω , the fraction of active managers in the j cohort (composed of managers employed for j periods) is equal to $\omega(1 - \omega)^j$. Denote by h_j the human capital of a manager who have been employed for j periods. Since human capital grows at the gross rate $g(\hat{\lambda}, \varepsilon)$, we have that $h_j = h_0 \Pi_{t=1}^j g(\hat{\lambda}, \varepsilon_t)$. Of course, this differs across managers of the same cohort because the growth rate is stochastic. The average human capital is then computed as

$$\bar{h} = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j h_j, \quad (46)$$

where E_j averages the human capital of all agents in the j -cohort. Because growth rates are serially independent, we have that $E_j h_j = h_0 E g(\hat{\lambda}, \varepsilon)^j$. Substituting in the above expression and solving we get

$$\bar{h} = \frac{h_0 \omega}{1 - (1 - \omega) E g(\hat{\lambda}, \varepsilon)}.$$

We now turn to the variance which is calculated as

$$\text{Var}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j (h_j - \bar{h})^2.$$

This can be rewritten as

$$\text{Var}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j (E_j h_j^2 - \bar{h}^2).$$

Using the serial independence of the growth rates, we have that $E_j h_j^2 = h_0^2 [E g(\hat{\lambda}, \varepsilon)^2]^j$. Substituting and solving we get

$$\text{Var}(h) = \frac{h_0^2 \omega}{1 - (1 - \omega) E g(\hat{\lambda}, \varepsilon)^2} - \bar{h}^2$$

To compute the inequality index we simply divide the variance by \bar{h}^2 , where \bar{h} is given by (46). This returns the inequality index (27).

To separate the *within* and *between* components of the inequality index, let's first rewrite the formula for the variance of h as follows:

$$\text{Var}(h) = \omega \sum_{j=0}^{\infty} (1 - \omega)^j \left[(E_j h_j^2 - \bar{h}_j^2) - (\bar{h}_j^2 - \bar{h}^2) \right],$$

where $\bar{h}_j = E_j h_j = h_0 Eg(\hat{\lambda}, \varepsilon)^j$ is the average human capital for the j cohort. Substituting the expression for h_j and \bar{h}_j and solving we get

$$\text{Var}(h) = \left(\frac{h^2 \omega}{1 - (1 - \omega) Eg(\hat{\lambda}, \varepsilon)^2} - \frac{h^2 \omega}{1 - (1 - \omega) (Eg(\hat{\lambda}, \varepsilon))^2} \right) + \left(\frac{h^2 \omega}{1 - (1 - \omega) (Eg(\hat{\lambda}, \varepsilon))^2} - \bar{h}^2 \right)$$

Dividing by \bar{h}^2 using the expression for \bar{h} derived in (46), we are able to write the inequality index as

$$\text{Inequality index} = \left(\frac{[1 - (1 - \omega) Eg(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1 - \omega) Eg(\hat{\lambda}, \varepsilon)^2]} - \frac{[1 - (1 - \omega) Eg(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1 - \omega) (Eg(\hat{\lambda}, \varepsilon))^2]} \right) + \left(\frac{[1 - (1 - \omega) Eg(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1 - \omega) (Eg(\hat{\lambda}, \varepsilon))^2]} - 1 \right) \quad (47)$$

The first term is the *within* cohorts inequality while the second term is the *between* cohorts inequality. Both terms are strictly increasing in $\hat{\lambda}$. ■

G Calibration

We use the 8 moments reported in the bottom section of Table 1 to calibrate 8 parameters. The mapping from the moments to the parameters is as follows:

- $\hat{\beta}$ is pinned down by the interest rate target, that is, $1/\hat{\beta} - 1 = 0.04$.
- ω is pinned down by the average life expectancy, that is, $1/\omega = 40$. Given the calibration of $\hat{\beta}$, in the model we use the discount factor $\beta = (1 - \omega)\hat{\beta} = 0.9375$.
- ψ is pinned down by the employment share in the financial sector together with the job finding rate in the sector, the probability ρ . Denote by S the number of workers employed in the nonfinancial sector and by U the number of workers with managerial ability, also employed in the nonfinancial sector. These workers flow into financial occupations at rate ρ , replacing financial managers who die at rate ω . Therefore, the number of financial managers evolves according to $1 - S_{t+1} = (1 - S_t)(1 - \omega) + U(1 - \omega)\rho_{t+1}$. The equivalent flow equation for workers with managerial ability is $U_{t+1} = U_t(1 - \omega)(1 - \rho) + \omega\psi$. After imposing steady state conditions, the two flow equations can be solved for

$$\psi = \frac{(\rho + \omega - \rho\omega)(1 - S)}{\rho(1 - \omega)},$$

where S has been determined by the employment share in the financial sector, ρ is a calibration target and ω has already been determined above.

- p is pinned down by the inequality index (coefficient of variation) in the financial sector. Section 3 has derived the inequality index in the financial sector as the square of the coefficient of variation in the cross sectional distribution of earnings. In the model with double-sided limited commitment the index can be derived analytically and takes the form

$$\text{Inequality index} = \frac{[1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)]^2}{\omega[1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)^2]} - 1.$$

The coefficient of variation is just the square root of this index. Because $\varepsilon \in \{0, 1\}$, we have that $Eg(\hat{\lambda}, \varepsilon) = 1 + p\hat{\lambda}$ and $Eg^2(\hat{\lambda}, \varepsilon) = 1 + 2p\hat{\lambda} + p^2\hat{\lambda}^2$. Therefore, the coefficient of variation is only a function of ω , $\hat{\lambda}$ and p . We can then use the calibrated value of ω and the targeted value of $\hat{\lambda}$ to pin down p .

- α is pinned down by the time spent innovating. In the model with double-sided limited commitment this maximizes the outside value of the manager and it is determined by the first order condition (17), that is, $\alpha/(1 - \hat{\lambda}) = \rho\beta\mathcal{B}p/(1 + \hat{\lambda})$.
- z is pinned down by the share of value added in the financial sector. First, in Section 3 we have derived the average human capital which is equal to

$$\bar{h} = h_0 \left[\frac{\omega}{1 - (1 - \omega)Eg(\hat{\lambda}, \varepsilon)} \right]$$

The output produced in the financial sector is $(1 - S)\bar{h}(1 - \hat{\lambda}^2)$ and the output produced in the nonfinancial sector is zh_0S . We can then determine z imposing that the output share of the financial sector is 8%.

- Finally, the parameters A and τ are pinned down by the probability of filling a vacancy and the probability of finding occupation in the financial sector. More specifically, we have $\rho = AX^{0.5}U^{-0.5}$ and $\phi = AX^{-0.5}U^{0.5}$. Given the calibration targets ρ and ϕ and the value of S determined above, we can use these two equations to solve for A . The free entry condition $\tau = \phi\bar{v}$ will then determine τ . Notice that, after imposing the targeted probabilities ρ and ϕ , we can solve for the steady state and, therefore, for the value of \bar{v} without the need of pre-setting the parameter τ . This parameter will then be determined residually without iteration.

References

- Albuquerque, Rui and Hugo A. Hopenhayn. 2004. "Optimal Lending Contracts and Firm Dynamics". *Review of Economic Studies*, 71(2), 285-315.
- Bell, Brian and John Van Reenen. 2010. "Banker's Pay and Extreme Wage Inequality in the UK," Centre for Economic Performance, LSE.
- Bolton, Patrick, Tanos Santos, and Jose Scheinkman. 2012. "Cream skimming in financial markets" SSRN 1770065.
- Cassis, Youssef. 2013. *Crisis & Opportunities: The Shaping of Modern Finance*, Cambridge University Press.
- Cheng, Ing-Haw, Harrison Hong, and Jose Scheinkman. 2012. "Yesterdays Heroes: Compensation and Creative Risk Taking," SSRN 1502762.
- Clementi, Gian Luca and Thomas F. Cooley. 2009. "Executive Compensation: Facts". NBER Working Papers No. 15426.
- Clementi, Gian Luca, Thomas F. Cooley, and Cheng Wang. 2006. "Stock Grants as a Commitment Device", *Journal of Economic Dynamics and Control*, 30(11), 2191-2216.
- Clementi, Gian Luca and Hugo Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121(1), 229-265.
- Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini. 2004. "Aggregate Consequences of Limited Contract Enforceability", *Journal of Political Economy*, 112(4), 817-847.
- Edmans, Alex and Xavier Gabaix. 2011. "The Effect of Risk on the CEO Market". *Review of Financial Studies*, 24, 2822-2863.
- Ellis, Charles D. 2008. *The Partnership: The Making of Goldman Sachs*. The Penguin Press, New York.
- Fishman, Michael and Peter M. DeMarzo. 2007. "Optimal Long-Term Financial Contracting", *Review of Financial Studies*, 20(6), 2079-2128.
- Grossman, Sanford and Oliver Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration". *Journal of Political Economy*, 94(4), 691-719.
- Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm". *Journal of Political Economics*, 98(6), 1119-58.
- Hosios, A. 1990. "On The Efficiency of Matching and Related Models of Search and Unemployment". *Review of Economic Studies*, 57(2), 279-298.

- Marcet, Albert, and Ramon Marimon. 2011. "Recursive Contracts". Economics Working Papers ECO2011/15, European University Institute.
- Marimon, Ramon and Vincenzo Quadrini. 2011. "Competition, Human Capital and Income Inequality with Limited Commitment". *Journal of Economic Theory*, 146(3), 976-1008.
- Philippon, Thomas, and Ariell Reshef. 2009. "Wages and Human Capital in the U.S. Financial Industry: 1909-2006". NBER Working Papers No. 14644.
- Phillipon, Thomas and Ariell Reshef. 2013 . "An International Look at the Growth of Modern Finance". *The Journal of Economic Perspectives*, 27(2),73-96.
- Quadrini, Vincenzo. 2004. "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard", *Journal of Monetary Economics*, 51(4), 713-751.
- Rajan, R. and Luigi Zingales. 2000. "The Governance of the New Enterprise", NBER WP 7958.
- Saez, Emmanuel and Thomas Piketty, "income Inequality in the United States - 1913-2002" *Quarterly Journal of Economics*. 118(1), 2003, 1-39.
- Shi, Shouyong. 2006. "Search Theory; Current Perspectives", *Palgrave Dictionary of Economics*.
- Spear, Stephen and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting", *Review of Economic Studies*, 54(4), 599-617.
- Wang, Cheng. 1997. "Incentives, CEO Compensation, and Shareholder Wealth in a Dynamic Agency Model. *Journal of Economic Theory*, 76(1), 72-105.