Risky Investments with Limited Commitment*

Thomas Cooley
Stern School of Business, New York University and NBER

Ramon Marimon
European University Institute, UPF - BarcelonaGSE, CEPR and NBER

Vincenzo Quadrini
University of Southern California and CEPR

February 28, 2014

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 and weakened the commitment between investors and managers. We show how increased competition and weaker commitment can raise the managerial incentives to undertake risky investment. In aggregate, this results in higher risk-taking, a larger and more productive financial sector, greater income inequality (within and across sectors), and a lower market valuation of financial institutions.

*We would like to thank Youssef Cassis, Robert Lucas, Facundo Piguillem, Kim Schoenholtz 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.

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 and more productive financial sector, higher compensation of financial executive and greater income inequality.

The increase in size and importance of the financial sector in the US economy is documented in Phillipon (2008) and Phillipon and Resheff (2009). Figure 1 shows that the GDP share of the financial industry doubled in size between 1970 and 2011. The share of employment has also increased but by less than the contribution to GDP. 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. Phillipon and Resheff (2013) show that the size of the financial sector has also increased in other countries.

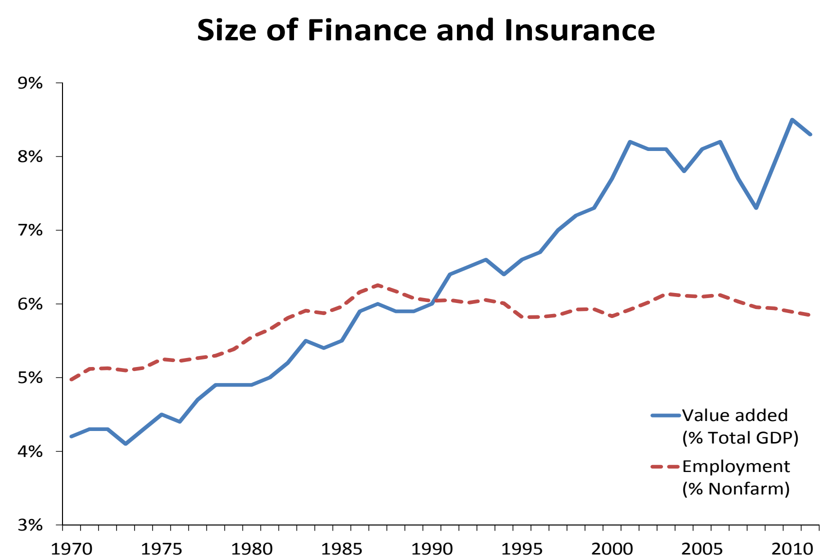

Figure 1: Share of Value Added and Employment

The increase in size was 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 the economy to nearly double. At the same time compensation of managers became more unequal in the financial sector. Figure 2 plots the income share of the top 5% of managerial positions in the financial sector compared to other occupations. As can be seen, the income concentration among managerial occupations has increased significantly compared to the rest of the economy.

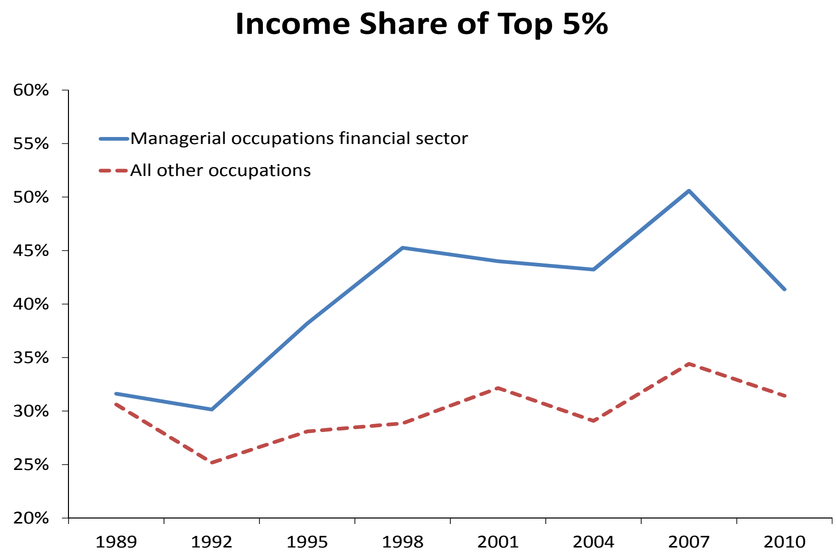

Figure 2: Share of the top 5%

Although productivity in the financial sector increased significantly, the market does not seem to value financial institutions when compared to the valuation of companies operating in other sectors of the economy. Figure 3 plots the average ratio of market to book value of equity for publicly listed financial and nonfinancial firms. Starting in the early 1980's, the market valuation of financial firms displays a flat trend while the valuation of nonfinancial firms has continued to grow. The fact that the market started to value financial firms less than other firms in the economy may be a reflection of compensation practices in a sector where managers retaining so much of the surplus. [1]

The changes described above took place during a period in which the organizational structure of the financial sector was also changing, with traditional partnerships replaced by public companies. Until 1970 the New York Stock Exchange prohibited member firms from being public companies. When the organizational restriction on financial companies was relaxed, 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.

[1] Since 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.

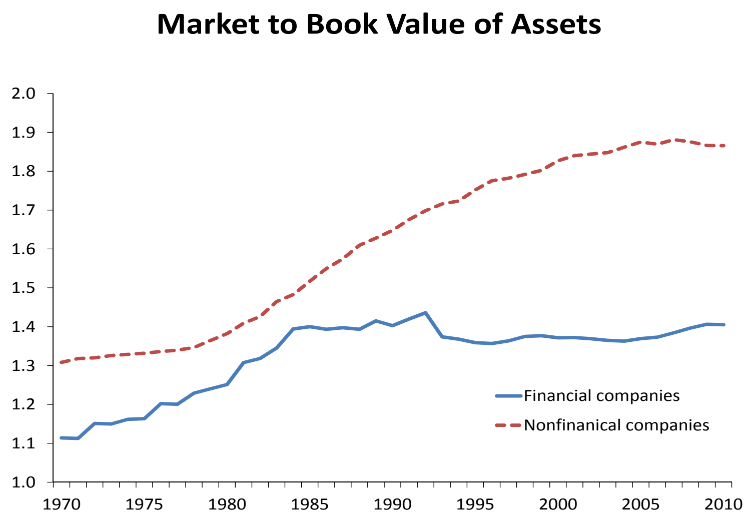

Figure 3: Average Market Value of Equity/Book Values of Equity

The transition from partnerships to public companies had two important implications. The first implication was to increase competition for managers in the financial sector. The second was to alter the structure of contractual arrangements between investors and managers in ways that weakened commitments.[2]

Why did the transformation from partnerships to corporations increased competition for managerial talent? As financial firms became public, they had greater access to capital (through the sales of shares) which facilitated their growth. But capital is only one of the production factors. Human capital is also important. Therefore, as more financial capital was coming in, more managerial capital was needed and this increased the competition (demand) for managers.[3]

Why did the transformation from partnerships to corporations weaken the commitment of investors and managers? Many argued that a partnership was a preferred form of organization for investment firms because managers and investors were the same people and it was the partners own assets that was at risk. Public companies, on the other hand, are organizational structures with significant separation between ownership

[2]The transition from partnership to public corporations also had some implications for the liability of partners and managers. However, in this paper, we do not address this particular aspect of the organizational change.

[3]Roy 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.

(shareholders) and investment control (managers), and it is well understood that they are characterized by significant agency problems.[4] Of course, in a world where contracts are fully enforceable, agency problems are solved with the optimal design of contracts. In reality, however, enforcement is limited and the transformation of partnerships to public companies has further reduced enforcement. On the one hand, shareholders could replace managers and renege future promises made to them. On the other, managers have some discretion in the operational decisions of the firm and could leave. The organizational change also increased the mobility of managers. In a partnership, the ownership shares were relatively illiquid so it was difficult for partners to liquidate their ownership positions and move to other firms. Also important was 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, represents a firm commitment to continued employment on the part of the other partners.

In this paper we focus on the these two implications of the organizational change: greater competition for managers and lower enforcement. We then ask whether they contributed to generate (*i*) greater risk-taking; (*ii*) a larger and more productive financial sector; (*iii*) higher compensation and greater income inequality (within and in relation to other sectors); (*iv*) lower stock market valuation of financial institutions.[5]

We address this question by developing a model in which investors compete for and hire managers to run investment projects, with each investor-manager pair representing a financial firm. Two features of the model are especially important. The first feature 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). The second important feature is that human capital can be transferred outside the firm by managers. This generates a conflict of interest between investors and managers: while the interest of investors is for the value of human capital inside the firm, managers also care about the outside value. As a result, the investment desired by investors may be smaller than the investment desired by managers. Then, if investors cannot control the firm policies either directly or indirectly through a credible compensation scheme, managers may

[4]This is largely consistent with the literature on incomplete contracts. 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.

[5]The New York Stock Exchange regulatory change mentioned above has been an important factor allowing financial corporations to become public companies. However, this does not tell us why they have chosen 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 goal of this paper is not to understand why financial companies have chosen to became public. Rather, we want to understand the consequences of having a financial industry changing from a partnership type of organization to public companies.

deviate from the optimal policies.

We first characterize the optimal contract with *one-sided* limited commitment. In this environment only the investor commits to the contract. We interpret this case as capturing the economic environment that prevailed in the period preceding the change in organizational structure (from partnerships to public companies). Although in this period there was not a clear separation between ownership and management, still, partners could quit the partnership, which motivates our choice of one-sided limited commitment to capture the contractual relationships of the earlier period. After studying the environment with one-sided limited commitment, we analyze the optimal contract with *double-sided* limited commitment. In this environment, contracts are not fully enforceable for both managers and investors. In particular, investors can reneg on promises made in the past and could replace the manager. We interpret this case as representative of the most recent period characterized by a clearer separation between ownership and management: When investors (shareholders) are different entities from managers, their commitment becomes relevant.

The main result of the paper shows that more competition for managerial talent has important implications for risk-taking and those implications depend on the contractual environment— i.e. one-sided vs. double-sided limited commitment. Risk taking is not exogenous in this model but depends on both competition and commitment. When investors commit, future payment promises are *credible* and they can be structured to deter managers from choosing riskier investments. As a result, higher competition induced by the organizational change from partnerships to public companies does not induce significant changes in risk taking per se. However, when investors do not commit to long-term contracts, promises of future payments are *not credible* and managers cannot be discouraged from choosing riskier investments. In this case a manager simply chooses the investment that maximizes her outside value, ignoring the cost that this imposes on the firm. As competition for managerial talent increases, so does the incentive to raise the outside value. Therefore, in the environment with double-sided limited commitment, risk taking rises with competition.

To make the outside value of managers endogenous and to study the implications for the whole economy, we embed this contractual micro structure in a general equilibrium model with two sectors—financial and nonfinancial. In the general model we formalize the increased competition for managers by lowering the cost to create jobs in the financial sector while the weakened commitment is captured by the shift to a regime with 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*) greater income inequality within and between sectors; and (*iv*) lower valuation of financial companies.

The organization of the paper is as follows. After relating the paper to the existing literature, Section 2 describes the theoretical model. Section 3 characterizes the optimal contract under different assumptions about commitment. Since the model is linear in human capital which grows over time, Section 4 reformulates the optimal contract with variables normalized by human capital. Section 5 embeds the micro structure in a general equilibrium whose properties are studied numerically in Section 6.1. Section 7 concludes.

1.1 Relation to the literature

The basic framework often used to study executive compensation is adapted from the principle-agent model of dynamic moral hazard by Spear and Srivastava (1987). Examples include Wang (1997), Quadrini (2004), Clementi and Hopenhayn (2006), Fishman and DeMarzo (2007). Albuquerque and Hopenhayn (2004) is also part of this class of models although the agency frictions are based on limited enforcement.

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. 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. In this paper we relax both assumptions: we endogenize the outside option of managers and we allow for the limited commitment of investors.[6]

The empirical facts described in the introduction have also motivated other studies. The models used in these studies can capture some of the empirical facts but we are not aware of models that can capture all of them simultaneously. We are also unaware of any study that connects change in the organizational structure of the financial sector with the increased competition for managerial talent. 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. In our paper, instead, we propose a model that could generate the empirical facts as a consequence of the organizational change that has taken place in the financial sector during the last three decades. In our model, the increase in risk is generated endogenously as a consequence of greater *competition* and weaker *commitment*. We show that when the level of risk is *endogenous*, it is optimal to lower risk in the constrained-efficient contract.

[6]Although 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.

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 framework in Section 5.

The financial sector is characterized by firms regulated by a contract between an investor—the owner of the firm—and a manager. In the case of a partnership we should think of the investor as the representative of all partners, who are also the managers of the firm. Effectively, each individual partner enters into a contractual relationship with all other partners who are represented by a fictitious 'investor'. In the case of public companies, instead, investors are distinct entities from managers. To simplify the analysis we assume that both a partnership and a public company is composed of a large number of partners so that the risk induced by the action of an individual partner or manager is negligible for the whole partnership or for the shareholders of the public company. Although this assumption may appear a major oversimplification, it allows us to capture some of the key differences between a partnership and a public company without loosing tractability. We will come back to this distinction and what this implies in term of optimal contracting.

We should think of managers as skilled workers who have the ability to run the firm and develop innovative projects. But managers could be mobile and when they choose to leave the firm, at least part of the know-how created with innovative projects can be transferred by them to other firms.

Preferences and technology. Preferences and technology are described without distinguishing the particular organizational structure (partnerships vs. public companies) and we will use the term 'investors' without specifying whether they are the representative of a partnership or the shareholders of a public company. The distinction will be made when we characterize the optimal contract since the different organizational structures imply different degrees of enforcement.

Investors (partners or shareholders) are modeled as risk-neutral agents who are the residual claimants to the output of the firm. Denoting by Y_t the output produced by the firm and C_t the manager's compensation, the expected lifetime utility of the investor is

$$V_0 = -C_0 + E_0 \sum_{t=1}^{\infty} \beta^t (Y_t - C_t) \equiv E_0 \sum_{t=0}^{\infty} \beta^t (\beta Y_{t+1} - C_t).$$

The reason production starts at at time 1 instead of time 0 is because production is realized with one period lag. Therefore, when a manager is hired at time zero, the manager will be paid C_0 but production starts at $t = 1$.

Managers are risk averse and they can choose time or effort allocated to two tasks: ordinary production activities and development of innovative projects. Denoting by ℓ_t the effort allocated to production activities and λ_t the effort allocated to innovation, the expected lifetime utility of a manager takes the form

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

where C_t is the manager's compensation (consumption). The period utility satisfies $u' > 0$, $u'' < 0$ and $e(0) = 0$, $e' > 0$, $e'' > 0$, $e'(0) = 0$, $e'(1) = \infty$. Both ℓ_t and λ_t are not directly observable by investors. These assumptions are relatively standard: the utility from consumption is strictly increasing and concave; the disutility from effort is strictly increasing and convex. Furthermore, the marginal disutility is zero initially but converges to infinity. This guarantees that the total effort $\ell_t + \lambda_t$ is always interior to the interval $[0, 1]$.

Managers are characterized by human capital h_t which is public information. The output produced by the firm in period $t + 1$, also publicly observable, is equal to

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

where the function $y(.)$ satisfies $y' > 0$, $y'' < 0$, $y'(0) = \infty$, $y'(1) = 0$, $y(0) = \underline{y}$. Therefore, output increases with the manager's human capital h_t and with time or effort allocated to production, ℓ_t. However, production effort displays decreasing returns. Furthermore, even if managers allocate all of their effort to innovation, there is still some production. Notice that production activities performed in period t generate output in period $t + 1$. This implies that, since ℓ_t is not observable by investors, they can only infer its value at time $t+1$ when they observe y_{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

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

where λ_t is the manager's effort allocated to innovation activities (which is not observed by investors) and $\varepsilon_{t+1} \in \{\underline{\varepsilon}, \bar{\varepsilon}\}$ is a publicly observed i.i.d. stochastic variable that can assume two possible values with probabilities $1 - p$ and p. The assumption that ε_{t+1} can take only two values is without loss of generality. Since h_t and ε_{t+1} are public information, the outcome of the innovative activity i_{t+1} is also publicly known at $t + 1$.

We think of i_{t+1} as the size of a new project that enhances the human capital of the manager only if the project is implemented in a firm—either the current or new firm. With its actual implementation, the human capital of the manager becomes $h_{t+1} = h_t + i_{t+1}$. If the new project i_{t+1} is not implemented in a firm—for instance, if the manager leaves the financial sector and finds occupation outside the financial industry—her human capital remains h_t. Therefore, if a project is implemented after the development stage, it becomes *embedded* human capital. Otherwise it fully depreciates. The importance of this assumption will become clear later.[7]

[7]The assumption that the pre-existing human capital does not depreciate when the manager leaves the

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 in a financial firm. The evolution of human capital can then be written as

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

Agency issues for managers. Managers have an option to quit and search for an offer in a new firm. If a manager chooses to quit, she will receive an offer with probability $\rho \in [0, 1]$. This probability captures the degree of *competition* for managers, that is, the ease with which a manager finds occupation in the financial sector after quitting the current employer. 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 $t + 1$ 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 manager's 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

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

For the moment we take the probability ρ and the outside value functions $\underline{Q}_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ as exogenous. At this stage we only assume that the $\underline{Q}_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ are strictly increasing and differentiable. However, when we extend the model to a general equilibrium in Section 5, the probability of an external offer and the outside value functions 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 effort ℓ_t and λ_t. Full control is allowed by the assumption that these two variables are not directly observable by investors. The investor can only infer the actual values of ℓ_t and λ_t in the next period after the realization of production, Y_{t+1}, and innovation output i_{t+1}. This implies that, in the absence of proper incentives, the values of ℓ_t and λ_t chosen by the manager may not be efficient. In particular, the manager may be tempted to increase λ_t and reduce ℓ_t in order to raise the outside value. Therefore, there are two sources of frictions in the decision problem of the manager: the ability to quit and the discretion in the choice of ℓ_t and 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 $\{C(\mathbf{H}^t, \mathbf{\Lambda}^{t-1})\}_{t=0}^{\infty}$ and effort*

financial industry is not essential for the qualitative properties of the model. It is only made to maintain the linear homogeneity in h_t. The alternative assumption that the whole human capital depreciates when the manager leaves the financial sector would lead to similar qualitative properties. However, we would lose the linear homogeneity property of the model in h_t. As we will see, this property allows us to work with a representative firm even if firms employ managers with different h_t.

decisions $\{\ell(\mathbf{H}^t, \mathbf{\Lambda}^{t-1})\}_{t=0}^{\infty}$ *and* $\{\lambda(\mathbf{H}^t, \mathbf{\Lambda}^{t-1})\}_{t=0}^{\infty}$*, conditional on the history of human capital* $\mathbf{H}^t = (h_0, \ldots, h_t)$ *and effort* $\mathbf{\Lambda}^{t-1} \equiv (\ell_0, \lambda_0, \ldots, \ell_{t-1}, \lambda_{t-1})$.

Notice that the payment made to the manager in period t is not conditional on ℓ_t and λ_t but only on past values. This is because ℓ_t and λ_t become public information only at $t + 1$.

Agency issues for investors: partnership vs. public companies. Agency issues could also emerge from the side of investors as they could renege on promises made to managers. The limited commitment of investors, however, depends on the particular organization structure. In a partnership, the investor is the representative of all partners. Therefore, it is unlikely that the investor reneges on promises made to the partners that they represent. Therefore, in characterizing the optimal contract in a partnership we assume that there is one-sided limited commitment: the representative of the partnership commits to the contract but individual managers do not commit.

In a public company, instead, investors are the shareholders of firms and they are distinct from managers. Because of this separation, the possibility of reneging on previous promises made to managers could become central to the contractual relationship between investors and managers of a public company. Of course, if we can write formal contracts in which future promises become legally binding, shareholders would not be able ex-post to renege on these promises. However, making the promises legally binding may not be desirable. Although not formalized in the model, making certain promises legally binding may discourage managerial effort, which is difficult to verify in a court. To capture this possibility we will assume that in a public company there is double-sided limited commitment: managers could quit and investors could renege on future promises.

3 Optimal contract

We start characterizing the optimal contract in a traditional partnership and then we move to the a public company. As argued above, an important difference between partnerships and public companies is the commitment of investors (the representative of the partners in a partnership and the shareholders in public companies). In both cases we make the simplifying assumption that in a firm—being a partnership or a public company—there is a large number of managers. In this way the risk faced by a firm as a consequence of the action taken by an individual manager is negligible. This allows us to solve the optimal contract as a relation between a risk-neutral investor and a risk-averse manager.

3.1 One-sided limited commitment: The case of partnerships

The optimal contract is 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 at the beginning of every period. This gives rise to the ***enforcement constraint***,

$$E_t \sum_{n=0}^{\infty} \beta^n \Big[u(C_{t+n}) - e(\ell_{t+n} + \lambda_{t+n})\Big] \geq D(h_{t-1}, h_t, \rho), \tag{4}$$

which must be satisfied for all $t \geq 1$. Notice that the contract starts at time zero but the constraint must be satisfied starting for $t \geq 1$. The participation constraint at time zero insures that the manager does not quit immediately after entering into the contractual relationship.

A second constraint takes into account that the manager has full control in the allocation of effort and could deviate from the ℓ_t and λ_t recommended by the planner (incentive-compatibility). Denote by $\hat{\ell}_t$ and $\hat{\lambda}_t$ the effort chosen by the manager when she deviates from the recommended effort. By deviating, the manager anticipates that she will leave the firm at the beginning of the next period. Therefore, $\hat{\ell}_t$ and $\hat{\lambda}_t$ are defined as

$$\Big\{\hat{\ell}_t, \hat{\lambda}_t\Big\} = \arg \max_{\ell,\lambda \in [0,1]} \Big\{u(C_t) - e(\ell + \lambda) + \beta E_t D\Big(h_t, g(\lambda, \varepsilon_{t+1})h_t, \rho\Big)\Big\}. \tag{5}$$

Since the manager quits the firm in the next period when she deviates from the recommended policies, the continuation value is the outside value $D\Big(h_t, g(\lambda, \varepsilon_{t+1})h_t, \rho\Big)$. It is important to point out that, the assumption that the manager quits at the beginning of next period after deviating is made to simplify the presentation but it is without loss of generality. In fact, the manager could still continue employment with the current firm after deviating. However, the continuation value received at $t + 1$ after deviating would still be $D\Big(h_t, g(\lambda, \varepsilon_{t+1})h_t, \rho\Big)$. This is because it is ex-ante optimal for the planner to impose the maximum punishment in case of deviation. Given the manager's option to quit, the maximum punishment is the value of quitting.

Since the manager would not get any benefit from producing once she anticipates leaving the firm, the optimal production effort is obviously $\hat{\ell}_t = 0$. The optimal innovation effort instead, will solve the first order condition

$$e_1(\hat{\lambda}_t) = \beta E_t D_2\Big(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho\Big) g_1(\hat{\lambda}_t, \varepsilon_{t+1})h_t. \tag{6}$$

We have used numerical subscripts to denote the derivative of a function with respect to a particular argument. Specifically, $D_2(.,.,.)$ denotes the derivative of the outside value with respect to the second argument and $e_1(.)$ is the derivative with respect to the first and only argument. The assumed properties of the function $e(.)$ guarantee an interior solution, that is, $\hat{\lambda}_t \in (0, 1)$. From now on, we will always denote with the hat sign the production and innovation efforts that maximizes the expected outside value net of dis-utility.

An important feature of the optimal deviations $\hat{\ell}_t$ and $\hat{\lambda}_t$ is that they are not affected by current compensation C_t since ℓ_t and λ_t are not observable by investors. Investors will infer the actual efforts at $t+1$ but at that point C_t has already been paid. The manager can still be punished at $t+1$ by cutting C_{t+1}. However, the ability to quit sets a lower bound to the feasible punishment. If C_t could be conditioned on $\hat{\ell}_t$ and $\hat{\lambda}_t$, investors could punish managers' deviation by reducing C_t. By further assuming that the utility function satisfies $u(0)=-\infty$, investors would have unlimited power to punish the mangers and, de-facto, they would not have discretion in the choice of ℓ_t and λ_t.

Another important feature is that, in the event of the deviation, the optimal production effort is zero, that is, $\hat{\ell}=0$. This is because production does not have any value for the manager when she anticipates leaving the firm. Given the substitutability between production effort and innovation effort in the utility function, when the manager deviates the marginal dis-utility from innovation declines. This may lead to higher innovation when the manager deviates from the recommended policy.

Given the optimal deviation $\hat{\ell}_t$ and $\hat{\lambda}_t$, the ***incentive-compatibility constraint*** at time t can be written as

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

The left-hand-side is the value that the manager receives if she chooses the effort recommended by the planner, ℓ_t and λ_t. The right-hand-side is the value achieved by deviating from the recommended policy, that is, when the manager chooses $\hat{\ell}_t=0$ and $\hat{\lambda}_t$ as determined in (5). As observed above, current compensation C_t cannot be contingent on the actual choice of ℓ_t and λ_t since these variables are not publicly observed at time t. Therefore, the current utility from consumption is the same with or without deviation.

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

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

The optimization problem is also subject to initial participation constraints for both the investor (the collective representative of all partners) and an individual partner which, for simplicity, we have omitted. These constraints only restrict the admissible values for the weight $\tilde{\mu}_0$.

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

$$\widetilde{W}(h,\tilde{\mu}) = \min_{\tilde{\chi},\tilde{\gamma}(\varepsilon')} \max_{C,\ell,\lambda} \Bigg\{ \beta y(\ell)h - C + \tilde{\mu}\Big(u(C) - e(\ell+\lambda)\Big) - \tilde{\chi}\Big(e(\ell+\lambda) - e(\hat{\lambda})\Big) + \beta E\Big[\widetilde{W}(h',\tilde{\mu}') - \Big(\tilde{\chi}+\tilde{\gamma}(\varepsilon')\Big)D(h,h',\rho)\Big]\Bigg\} \quad (9)$$

$$\textbf{s.t.} \quad h' = g(\lambda,\varepsilon')h, \quad \tilde{\mu}' = \tilde{\mu} + \tilde{\chi} + \tilde{\gamma}(\varepsilon'),$$

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

The variable $\tilde{\mu}_t$, which becomes a state variable for the recursive formulation of the planner's problem, captures the value of the contract for the manager. This variable evolves over time according to $\tilde{\mu}' = \tilde{\mu} + \tilde{\chi} + \tilde{\gamma}(\varepsilon')$. Therefore, any time the incentive-compatibility constraint or the enforcement constraint are binding, the manager's value increases. This translates to higher promises that are necessary to prevent the manager from deviating and quitting.

Optimal partnership policies. Differentiating problem (9) with respect to the manager's consumption C we obtain,

$$C_t = u_1^{-1}\left(\frac{1}{\tilde{\mu}_t}\right), \quad (10)$$

which characterizes the ***compensation policy*** as a function of the state variable $\tilde{\mu}_t$.

As we have seen, this variable evolves according to the law of motion $\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})$. Therefore, anytime the enforcement and/or the incentive-compatibility constraints are binding, the manager's consumption increases. Since λ_t is always positive, h_t grows in expectation and with it the outside value of the manager $D(h_t, h_{t+1}, \rho)$. This implies that the enforcement constraint becomes binding at some point in the future and raises $\tilde{\mu}$. From equation (10) we can then see that the growth in $\tilde{\mu}$ is inherited by consumption. Therefore, the optimal partnership contract does not provide full insurance.

The ***production policy*** is characterized by the first-order condition,

$$(\tilde{\mu}_t + \tilde{\chi}_t)e_1(\ell_t + \lambda_t) = \beta y_1(\ell_t)h_t. \quad (11)$$

The left-hand-side is the marginal cost of production in terms of effort dis-utility. The cost increases when the incentive-compatibility constraint is binding. This is because, to insure that the manager does not deviate from the recommended policy, her utility has to increase. One way to increase utility is by reducing effort. The right-hand-side of (11) is the marginal benefit of production, that is, the increase in output.

The ***innovation policy*** is characterized by the first-order condition with respect to λ. Using $g(\lambda, \varepsilon') = 1 + \lambda\varepsilon'$, the optimality condition can be written as

$$(\tilde{\mu}_t + \tilde{\chi}_t)\, e_1(\ell_t + \lambda_t) = \beta E_t\Big[W_1\left(h_{t+1}, \tilde{\mu}_{t+1}\right) - \Big(\tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})\Big) D_2\left(h_t, h_{t+1}, \rho\right)\Big] h_t\varepsilon_{t+1}. \quad (12)$$

The left-hand side is the marginal cost of innovation per unit of human capital, which is increasing in ℓ_t, λ_t, $\tilde{\mu}_t$ and $\tilde{\chi}_t$. The right-hand-side is the expected marginal benefit from investing, net of participation costs. Combining (11) and (12) we obtain

$$y_1(\ell_t) = E_t\Big[W_1\left(h_{t+1}, \tilde{\mu}_{t+1}\right) - \Big(\tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})\Big) D_2\left(h_t, h_{t+1}, \rho\right)\Big]\varepsilon_{t+1}. \quad (13)$$

Binding incentive-compatibility and enforcement constraints imply positive values of $\tilde{\chi}_t$ and $\tilde{\gamma}_t(\varepsilon_{t+1})$ which tend to reduce the right-hand-side of this equation. Therefore, to the extent that the right-hand-side term declines with $\tilde{\chi}_t$ and $\tilde{\gamma}_t(\varepsilon_{t+1})$, we have that binding constraints induce an increase in production effort ℓ_t (since the marginal productivity of effort is decreasing. Then, going back to condition (11), we can see that the innovation effort has to decline.

Intuitively, to retain the manager, the value of staying must increase or the value of quitting must decline. The value of staying can be increased by promising higher compensation and by requiring lower effort (that is, by changing λ_t). The value of quitting can be reduced by choosing a lower λ_t. Therefore, λ_t decreases when the constraints become binding. However, as λ_t declines, the marginal dis-utility from production effort falls. It then becomes optimal to allocate more effort to production. Effectively, it becomes optimal to substitute innovation effort with production effort. The the reduction in innovation effort, however, is bigger than the increase in production effort.

We are particularly interested in understanding how higher competition (captured by a higher value of ρ) affects 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). We obtain the following result (formally proved in Appendix B).

Proposition 1 *If $W_{1,2}(h_{t+1}, \tilde{\mu}_{t+1}) \leq 0$, more competition for managers (higher ρ) results in lower innovation λ_t when the enforcement and incentive-compatibility constraints are binding.*

As we discuss in Appendix B, the condition $W_{1,2}(h_{t+1}, \tilde{\mu}_{t+1}) \leq 0$ is fairly general. In particular, this condition is satisfied when the manager's utility from consumption takes the logarithmic form as we will see in Section 4.

3.2 Double-sided limited commitment: The case of public companies

In the environment with double-sided limited commitment, which we think represents the contractual environment in public companies, managers are free to leave the company and investors can renege promises made to managers. This implies that investors renegotiate whenever the value of the contract for the manager exceeds the outside value. As a result, the planner also faces the constraint that the value of the contract for the manager cannot exceed the outside value of the manager.

The limited commitment of the investor alters the optimization problem (9) in several dimensions. First, in anticipation of investor's renegotiation, the manager always chooses the allocation of effort that maximizes the outside value. Therefore, with double-sided limited commitment we have that $\ell_t = \hat{\ell}_t$ and $\lambda_t = \hat{\lambda}_t$. This also implies that the incentive-compatibility constraint is no longer relevant and $\tilde{\chi}$ can be set to zero.

The second modification is that the variable $\tilde{\mu}_{t+1}$, the weight assigned by the planner to the manager in the next period, is no longer dependent on $\tilde{\mu}_t$. The dependence of $\tilde{\mu}_{t+1}$ from $\tilde{\mu}_t$ (through 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 in the next period. Therefore, even if the enforcement constraint is not binding tomorrow, the new weight assigned to the manager will not be reduced. Without commitment, however, promises made today and captured by the variable $\tilde{\mu}_t$, are no longer relevant. Therefore, $\tilde{\mu}_{t+1}$ is exclusively determined by the multiplier associated with the enforcement constraint in the next period, that is, $\tilde{\mu}_{t+1} = \tilde{\gamma}_t(\varepsilon_{t+1})$. The contractual problem can be written as

$$W(h, \tilde{\mu}) = \min_{\tilde{\gamma}(\varepsilon')} \max_{C} \left\{ \beta y(0)h - C + \tilde{\mu}\Big(u(C) - e(\hat{\lambda})\Big) + \beta E\Big[W\Big(h', \tilde{\mu}'\Big) - \tilde{\gamma}(\varepsilon')D\Big(h, h', \rho\Big)\Big]\right\} \tag{14}$$

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

The contract simply prescribes a consumption plan which is determined by (10) with $\tilde{\mu}' = \tilde{\gamma}(\varepsilon')$. The production effort is $\hat{\ell} = 0$ and the innovation effort solves the first order condition (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 innovation for the manager. This is stated formally in the next proposition.

Proposition 2 *With double-sided limited commitment a higher ρ results in higher innovation $\hat{\lambda}$.*

This is a key result of this analysis. Together, Propositions 1 and 2 show that the effect of more competition for managers on risk-taking depends crucially on whether investors commit to the contract. Higher competition increases risk-taking only when there is limited commitment of both investors and managers. To the extent that the organizational change from partnerships to public companies increased mobility (higher ρ) and weakened commitment (especially for investors), we should observe higher risk-taking.

4 Normalization with log-utility

Since human capital grows on average over time, so does the value of the contract for both the manager and the investor. It is then convenient to normalize the growing variables so that we can work with a stationary formulation of the contracting problem. This is especially convenient when the utility of managers and the outside values take the logarithmic form.

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

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

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 in the extension to a general equilibrium that with log utility they take exactly these forms.

We start by normalizing the value of the contract for the investor which can be expressed recursively as $V_t = \beta y(\ell_t)h_t - C_t + \beta E_t V_{t+1}$. This can be rewritten as,

$$v_t = \beta y(\ell_t) - c_t + \beta E_t g(\lambda_t, \varepsilon_{t+1}) v_{t+1}, \tag{15}$$

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 $Q_t = \ln(C_t) - e(\ell_t + \lambda_t) + \beta E_t Q_{t+1}$ which can be rewritten After defining $q_t = Q_t - \mathcal{B}\ln(h_t)$, we can rewrite it in normalized form as,

$$q_t = \ln(c_t) - e(\ell_t + \lambda_t) + \beta E_t \Big[\mathcal{B} \ln\Big(g(\lambda_t, \varepsilon_{t+1})\Big) + q_{t+1} \Big]. \tag{16}$$

Next we consider the enforcement constraint after the realization of ε_{t+1},

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

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 rewritten as

$$q_{t+1} \geq (1-\rho)\underline{q} + \rho\overline{q} - (1-\rho)\mathcal{B}\ln\Big(g(\lambda_t, \varepsilon_{t+1})\Big). \tag{17}$$

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 last term in (17) would disappear. The value of quitting would still depend on ρ but it would not affect the choice of λ_t.

The constraint that insures that the manager chooses the optimal allocation of effort (incentive-compatibility) is,

$$-e(\ell_t + \lambda_t) + \beta E_t Q_{t+1}\Big(g(\lambda_t, \varepsilon_{t+1})h_t\Big) \geq$$
$$-e(\hat{\lambda}_t) + \beta E_t\Big[(1-\rho)\cdot \underline{Q}_{t+1}(h_t) + \rho \cdot \overline{Q}_{t+1}\Big(g(\hat{\lambda}_t, \varepsilon_{t+1})h_t\Big)\Big],$$

where λ_t is the investment recommended by the contract and $\hat{\lambda}_t$ is the investment chosen by the manager under the assumption that she will quit at the beginning of next period. After normalizing, the incentive-compatibility constraint becomes,

$$-e(\ell_t + \lambda_t) + \beta E_t\Big[q_{t+1} + \mathcal{B}\ln\Big(g(\lambda_t, \varepsilon_{t+1})\Big)\Big] \geq$$
$$-e(\hat{\lambda}_t) + \beta E_t\Big[(1-\rho)\underline{q} + \rho\bar{q} + \rho\mathcal{B}\ln\Big(g(\hat{\lambda}_t, \varepsilon_{t+1})\Big)\Big]. \quad (18)$$

We can now provide a more explicit characterization of the manager's optimal deviation $\hat{\lambda}_t$. Using $g(\lambda, \varepsilon) = 1 + \lambda\varepsilon$, the optimal deviation condition (6) can be written,

$$e_1(\hat{\lambda}_t) = \rho\beta\mathcal{B}E_t\left(\frac{\varepsilon_{t+1}}{1 + \hat{\lambda}_t\varepsilon_{t+1}}\right), \quad (19)$$

We can now see more explicitly that, given the properties of the dis-utility function $e(.)$, $\hat{\lambda}$ increases in the probability ρ, as stated more generally in Proposition 2. Therefore, when the manager faces better outside options, the strategic incentive to innovate increases.

One-sided limited commitment: The case of partnerships. 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,

$$v(q) = \max_{\lambda, c, q(\varepsilon')}\Big\{\beta y(\ell) - c + \beta E g(\lambda, \varepsilon')v\Big(q(\varepsilon')\Big)\Big\} \quad (20)$$

subject to (16), (17), (18).

The solution provides the effort policies $\ell = \varphi^\ell(q)$ and $\lambda = \varphi^\lambda(q)$, 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 the innovation policy λ and the initial human capital h, we can reconstruct the whole sequence of human capital through the law of motion $h' = g(\lambda, \varepsilon')$. Then we can reconstruct the original, non-normalized variables $C = ch$, $Q = q + \mathcal{B}\ln(h)$ and $V = vh$.

There is a one-to-one mapping from the normalized policies to the original (non-normalized) variables. To characterize the optimal contract we can focus on the normalized policies which satisfy the first order conditions

$$c = \mu, \tag{21}$$

$$(\mu + \chi)e_1(\ell + \lambda) = \beta y_1(\ell), \tag{22}$$

$$(\mu + \chi)e_1(\ell + \lambda) = \beta E\left[v\Big(q(\varepsilon')\Big) + \frac{\mathcal{B}[\mu + \chi + (1-\rho)\gamma(\varepsilon')]}{1 + \lambda\varepsilon'}\right], \tag{23}$$

$$\mu(\varepsilon') = \frac{\mu + \chi + \gamma(\varepsilon')}{1 + \lambda\varepsilon'}. \tag{24}$$

The variables μ, $\gamma(\varepsilon')$ and χ are the Lagrange multipliers for constraints (16)-(18). These multipliers are related to the multipliers used in Section 3 as follows: $\mu = \tilde{\mu}/h$, $\tilde{\gamma}(\varepsilon')/h$ and $\tilde{\chi}/h$. The detailed derivation is provided in Appendix C.

Double-sided limited commitment: The case of public companies. With double-sided limited commitment, investors renegotiate promises that exceed the outside value of the manager. Therefore, the value of the contract for the manager is always equal to the outside value, that is, the enforcement constraint is always satisfied with equality. Anticipating renegotiation, the best strategy for the manager is to choose production effort $\hat{\ell} = 0$ and innovation effort $\hat{\lambda}$ as determined by condition (19). Problem (14) can then be reformulated in normalized form as,

$$v(q) = \max_{c, q(\varepsilon)} \left\{ \beta y(0) - c + \beta E g(\hat{\lambda}, \varepsilon) v\Big(q(\varepsilon)\Big) \right\} \tag{25}$$

subject to

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

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

Problem (25) is a special case of problem (20) where we have replaced the incentive-compatibility constraint (18) with $\ell = \hat{\ell} = 0$ and $\lambda = \hat{\lambda}$, and the enforcement constraint (17) is always binding. Notice that the decision variables c and $q(\varepsilon)$ are fully determined by the promise-keeping and enforcement constraints. Therefore, the problem can be solved without performing any optimization, besides solving for $\hat{\lambda}$.

4.1 Contract properties

In this subsection we illustrate the properties of the optimal contract numerically. The specific parameter values will be described in Section 6.1 where we conduct a quantitative analysis with the general model. The computational procedure used to solve the optimal contract is described in Appendix E[8].

As we have seen, the solution to the contractual problem (20) 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 (21)-(24). This is also the case for the solution to problem (25) in the environment with double-sided limited commitment.

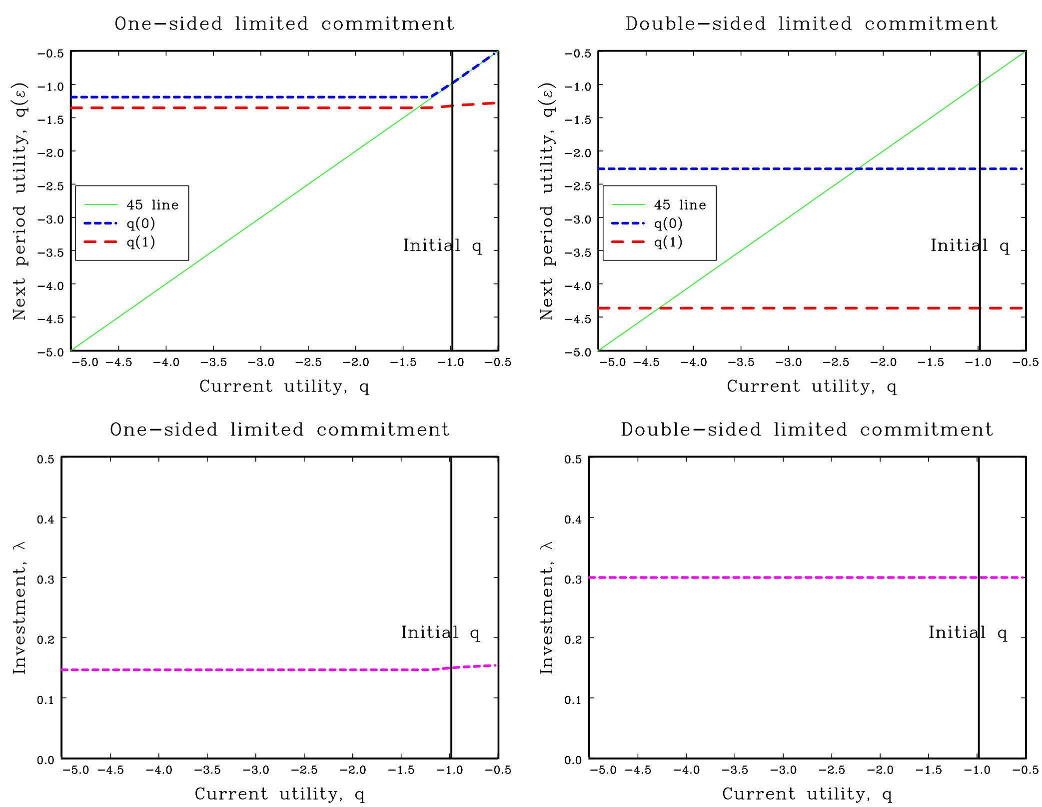

Figure 4: Continuation utilities and investment with one-sided and double-sided limited commitment.

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

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 ($\varepsilon = 1$), 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 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.

5 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, $H = h_0 S$. 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

$$\underline{Q}_t(h) = \ln(h) + \beta\Big[(1 - \rho_{t+1}) \cdot \underline{Q}_{t+1}(h) + \rho_{t+1} \cdot \overline{Q}_{t+1}(h)\Big]. \tag{26}$$

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.

5.1 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 $\overline{q}_t$ is decided, the investor offers $\overline{Q}_t = \overline{q}_t + \mathcal{B}\ln(h)$ 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

$$\phi_t(\bar{q}_t)\, v_t(\bar{q}_t) = \tau. \tag{27}$$

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 (26). This can be rewritten as

$$\underline{q}_t = \ln(1) + \beta\Big[(1-\rho_{t+1})\cdot \underline{q}_{t+1} + \rho_{t+1}\cdot \overline{q}_{t+1}\Big]. \tag{28}$$

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 (27), (28) and (39);*
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* $\overline{q}$ *offered to the manager; i.e.* $\overline{q}'(\rho) > 0$.

5.2 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

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

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

$$\text{Ave}(h) = h_0\left[\frac{\omega}{1-(1-\omega)Eg(\hat{\lambda},\varepsilon)}\right], \tag{29}$$

$$\text{Inequality index} = \frac{[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)]^2}{\omega[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)^2]}-1. \tag{30}$$

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:

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

where $\mathrm{Ave}_j(h)$ 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 $\hat{\lambda}$.

6 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.

6.1 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.

Table 1: Parameters and calibration moments.

Parameters

$\hat{\beta}$	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

Calibration moments

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

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 enforcement 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.

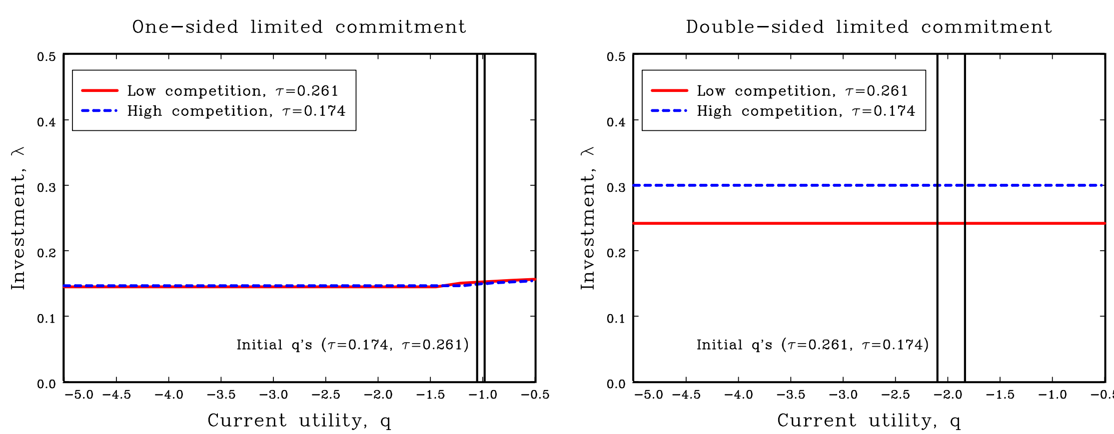

Figure 5: Steady state investment policies for different τ in the environments with one-sided and double-sided limited commitment.

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 increase 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.

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

	One-sided limited commitment	*Double-sided limited commitment*
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

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.[9] 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 maybe tempted to recover his *ex-post* relative losses due to increased competition by reneging on his commitments. Such a move to a double-sided limited commitment economy may reduce 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.

7 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 investors[10]. 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

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

[10]See Footnotes ?? and ?? for a brief reference to alternative explanations.

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 sectors and patents on financial instrument are *rare avis* and difficult to enforce).[11]

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

[11] Although 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).

previous organizational form, that is, the traditional partnership.[12]

[12] "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 Partnesrship form does not fit easily into the contracting structure that we use to characterize incentives in modern fiancial 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:

$$VP(h_t) = \max_{\lambda_t} \{\beta u(y(\lambda_t)h_t) - e(\lambda_t) + \beta E_t VP(h_{t+1})\} \tag{31}$$

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

which results in the following Euler's equation:

$$\beta^2 E_t u'(C_{t+2}) y(\lambda^*_{t+1}) g_\lambda(\lambda^*_t, \varepsilon_{t+1}) h_t \le -\beta u'(C_{t+1}) y'(\lambda^*_t) h_t + e'(\lambda^*_t), \tag{32}$$

with equality if $\lambda^*_t > 0$.

If we assume that the partner has log utility preferences for consumption, (32) simplifies to:

$$p\beta^2 (1+\lambda^*_t)^{-1} \le -\beta \frac{y'(\lambda^*_t)}{y(\lambda^*_t)} + e'(\lambda^*_t) \tag{33}$$

If, in addition, we use the functional forms $e(\lambda) = -\alpha(1-\lambda)$ and $y(\lambda) = 1-\lambda^2$, then (33) simplifies to

$$p\beta^2 \le f(\lambda) \equiv \frac{\alpha + \lambda(2+\alpha)}{1-\lambda}. \tag{34}$$

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:

$$\min_{\chi,\gamma(\varepsilon')} \max_{c,\lambda} \quad \Big\{ \beta y(\lambda) - c + \mu\Big(u(ch) - e(\lambda)\Big) - \chi\Big(e(\lambda) - e(\hat{\lambda})\Big)$$

$$+\beta E\Big[v(\mu')g(\lambda, \varepsilon') + \left(\mu + \chi + \gamma(\varepsilon')\right) Q(h', \mu')$$

$$-\chi D\Big(h, g(\hat{\lambda}, \varepsilon')h, \rho\Big) - \gamma(\varepsilon')D(h, h', \rho)\Big]\Big\}$$

$$\textbf{s.t.} \quad h' = g(\lambda, \varepsilon')h, \quad \mu' = \left(\mu + \chi + \gamma(\varepsilon')\right)/g(\lambda, \varepsilon'),$$

and the corresponding first-order condition with respect to λ is given by (12):

$$(\mu + \chi)\, e_\lambda(\lambda) - \beta y_\lambda(\lambda) \quad \geq \quad \beta E\Big[\Big(v\left(\mu'\right) + \left(\mu + \chi + \gamma(\varepsilon')\right) Q_h(h', \mu')$$

$$-\gamma(\varepsilon')D_2\left(h, h', \rho\right)\Big)\varepsilon'\Big].$$

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 $D_{2,3}\left(h, h', \rho\right)$. By the definition of D, (3), $D_{2,3} > 0$ and, therefore, this direct effect of the enforcement constraint makes investment more costly. Furhtermore, 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 $\beta E\Big[Q_h(h', \mu')\varepsilon'\Big] - e_\lambda(\lambda) - \beta y_\lambda(\lambda) \leq 0$, where the inequality follows from the fact that otherwise $\chi = 0$; similarly, increasing $\gamma(\varepsilon')$ results in $Q_h(h', \mu') - D_2\left(h, h', \rho\right) \leq 0$. There only remain the effects of increasing μ', which are given by $v'\left(\mu'\right) < 0$ 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,

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

and, therefore,

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

In sum, $u_{h,\mu}\left(h\mu\right) \leq 0$ 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 $Q_{h,\mu}\left(h\mu\right) > 0$. 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. $Q_{h,\mu}\left(h\mu\right) = 0$.

C The first-order conditions of Problem (20)

Let μ, $\gamma(\varepsilon)$ and χ be the lagrange multipliers associated with the promise-keeping, enforcement and incentive-compatibility constraints. The lagrangian can be written as

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

The terms $\bar{d}$ and $d(\hat{\lambda})$ collect variables and functions that are not affected by the contract policies λ, c and $q(\varepsilon)$. The first order conditions with respect to these three variables are, respectively,

$$-\beta - (\mu+\chi)e_1(\lambda) + \beta\sum_{\varepsilon}\Bigg[g_1(\lambda,\varepsilon)v\Big(q(\varepsilon)\Big) + \mathcal{B}\left(\frac{g_1(\lambda,\varepsilon)}{g(\lambda,\varepsilon)}\right)\Big(\mu+\chi+(1-\rho)\gamma(\varepsilon)\Big)\Bigg]p(\varepsilon) = 0$$

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

$$g(\lambda,\varepsilon)v_1\Big(q(\varepsilon)\Big) + \Big(\mu+\chi+\gamma(\varepsilon)\Big) = 0$$

Substituting the envelope condition $v_1(q) = -\mu$ and using the functional forms of $y(\lambda)$ and $g(\lambda, \varepsilon)$ we obtain equations (23)-(24).

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 $\overline{Q}_t^*(h)$ is the value of the equilibrium contract, then for any $\overline{Q}_t(h)$ in a neighbourhood $\mathcal{N}(\overline{Q}^*)$ of $\overline{Q}_t^*(h)$, the following condition is satisfied,

$$\rho_t\Big(h, \overline{Q}_t(h)\Big) \cdot \Big[\overline{Q}_t(h) - \underline{Q}_t(h)\Big] = \rho_t\Big(h, \overline{Q}_t^*(h)\Big) \cdot \Big[\overline{Q}_t^*(h) - \underline{Q}_t(h)\Big], \tag{35}$$

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 $\overline{Q}_t^*(h)$ as given and choose the contract by solving the problem

$$\max_{\overline{Q}_t(h)} \quad \Big\{\phi_t\Big(h, \overline{Q}_t(h)\Big) \cdot V\Big(h, \overline{Q}_t(h)\Big)\Big\} \tag{36}$$

$$\text{subject to (35),}$$

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 investor[13] $V_t(h, \overline{Q}_t(h))$.

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

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

$$\phi_t\Big(h, \overline{Q}_t(h)\Big) \cdot V_t\Big(h, \overline{Q}_t(h)\Big) = \tau h. \tag{37}$$

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,

[13] Given 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'.

that is, $V_t\Big(h, Q_t(h)\Big) = v_t(q_t)h_t$. Therefore, the free entry condition can be rewritten in normalized form as

$$\phi_t(\overline{q}_t) \cdot v_t(\bar{q}_t) = \tau. \tag{38}$$

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)[14].

The investor's problem (36) can be rewritten as

$$\begin{aligned} \bar{q}_t &= \arg\max_{q} \; \Big\{\phi_t(q) \cdot v_t(q)\Big\} \\ & \quad \text{subject to} \\ & \quad \rho_t(q)(q - \underline{q}_t) = \rho_t(\bar{q}_t^*)(\bar{q}_t^* - \underline{q}_t), \; \forall q \in \mathcal{N}(\bar{q}_t^*) \end{aligned}$$

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

$$1 - \eta = \frac{-v_t'(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}{v_t(\bar{q}_t) - v_t'(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}. \tag{39}$$

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. $v_\epsilon(\bar{q}) \equiv -\frac{v'(\bar{q})\bar{q}}{v(\bar{q})}$. 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 (39) can be written as

$$\frac{1-\eta}{\eta} = v_\epsilon(\bar{q})\frac{\bar{q} - \underline{q}}{\bar{q}}. \tag{40}$$

In a stationary equilibrium, using (28) we obtain:

[14] In 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.

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

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 (20) 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

$$\beta\left[v\Big(q(1)\Big)+\left(\frac{\mathcal{B}}{1+\lambda}\right)\Big(\mu+\chi+(1-\rho)\gamma(1)\Big)\right]p=-\beta y_{\lambda}(\lambda)+\frac{\alpha(\gamma+\chi)}{1-\lambda} \tag{41}$$

$$c=\gamma \tag{42}$$

$$g(\lambda,\varepsilon)\psi\Big(q(\varepsilon)\Big)=\mu+\chi+\gamma(\varepsilon) \tag{43}$$

$$v=\beta y(\lambda)-c+\beta\sum_{\varepsilon}g(\lambda,\varepsilon)\Psi\Big(q(\varepsilon)\Big)p(\varepsilon) \tag{44}$$

$$q=\ln(c)+\alpha\ln(1-\lambda)+\beta\sum_{\varepsilon}\Big(\mathcal{B}\ln\Big(g(\lambda,\varepsilon)\Big)+q(\varepsilon)\Big)p(\varepsilon) \tag{45}$$

$$\chi\left\{\alpha\ln(1-\lambda)+\beta\sum_{\varepsilon}\left[q(\varepsilon)+\mathcal{B}\ln\Big(g(\lambda,\varepsilon)\Big)\right]p(\varepsilon)\right.$$
$$\left.-\alpha\ln(1-\hat{\lambda})-\beta\sum_{\varepsilon}\left[(1-\rho)\underline{q}+\rho\bar{q}+\rho\mathcal{B}\ln\Big(g(\hat{\lambda},\varepsilon)\Big)\right]p(\varepsilon)\right\}=0 \tag{46}$$

$$\gamma(\varepsilon)\left[q(\varepsilon)-(1-\rho)\underline{q}-\rho\bar{q}+(1-\rho)\mathcal{B}\ln\Big(g(\lambda,\varepsilon)\Big)\right]=0 \tag{47}$$

The first three equations are the first order conditions with respect to λ, c, $q(\varepsilon)$, respectively. Equation (44) defines the value for the investor and equation (45) is the promise-keeping constraint. Equations (46) and (47) formalize the Kuhn-Tucker conditions for the incentive-compatibility and enforcement constraints.

Notice that equations (46) and (47) 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 (28). 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 (38) and the bargaining condition (39). 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 mangers of the same cohort because the growth rate is stochastic. The average human capital is then computed as

$$\bar{h} = \omega \sum_{j=0}^{\infty} (1-\omega)^j E_j h_j, \tag{48}$$

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

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

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

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

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

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

To compute the inequality index we simply divide the variance by $\bar{h}^2$, where $\bar{h}$ is given by (48). This returns the inequality index (29).

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

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

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

$$\begin{aligned}\mathrm{Var}(h) &= \left(\frac{h^2\omega}{1-(1-\omega)Eg(\hat{\lambda},\varepsilon)^2} - \frac{h^2\omega}{1-(1-\omega)(Eg(\hat{\lambda},\varepsilon))^2}\right) + \\ &\quad \left(\frac{h^2\omega}{1-(1-\omega)(Eg(\hat{\lambda},\varepsilon))^2} - \bar{h}^2\right)\end{aligned}$$

Dividing by $\bar{h}^2$ using the expression for $\bar{h}$ derived in (48), we are able to write the inequality index as

$$\begin{aligned}\text{Inequality index} &= \left(\frac{[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)]^2}{\omega[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)^2]} - \frac{[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)]^2}{\omega[1-(1-\omega)(Eg(\hat{\lambda},\varepsilon))^2]}\right) + \\ &\quad \left(\frac{[1-(1-\omega)Eg(\hat{\lambda},\varepsilon)]^2}{\omega[1-(1-\omega)(Eg(\hat{\lambda},\varepsilon))^2]} - 1\right) \qquad (49)\end{aligned}$$

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 5 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 (19), 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 5 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 skimmming in fiancial 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, Emmanuael 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), 72Â–105.