

Risky Investments with Limited Commitment*

Thomas Cooley[†] Ramon Marimon[‡] Vincenzo Quadrini[§]

June 12, 2013

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 with 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 but also weakened the commitment between investors and managers. We show that the increased competition and the weaker commitment has raised the managerial incentives to undertake risky investment. In the general equilibrium, this change results in a larger financial sector, higher risk-taking and greater income inequality.

1 Introduction

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

The increase in the size and importance of the financial sector in the United States has been documented by Phillipon (2008) and Phillipon and Resheff (2009). This is also shown in Figure 1 which plots the shares of the financial industry in value added and employment since the late 1940's. The contribution of the finance industry to GDP doubled in size between 1970 and 2006. The share of employment has also increased but by less than the contribution to value added. This is especially noticeable starting in

*We would like to thank Facundo Piguillem and other seminar participants at the Banca d'Italia, European University Institute, Banque de France, Tel Aviv University and Hebrew University of Jerusalem for their comments on previous versions of this work.

[†]Stern School of Business, New York University and NBER

[‡]European University Institute, UPF - BarcelonaGSE, CEPR and NBER

[§]University of Southern California and CEPR

the mid 1980s when the share of employment stopped growing while the share of value added continued to expand. Accordingly, we observe a significant increase in productivity compared to the remaining sectors of the economy.

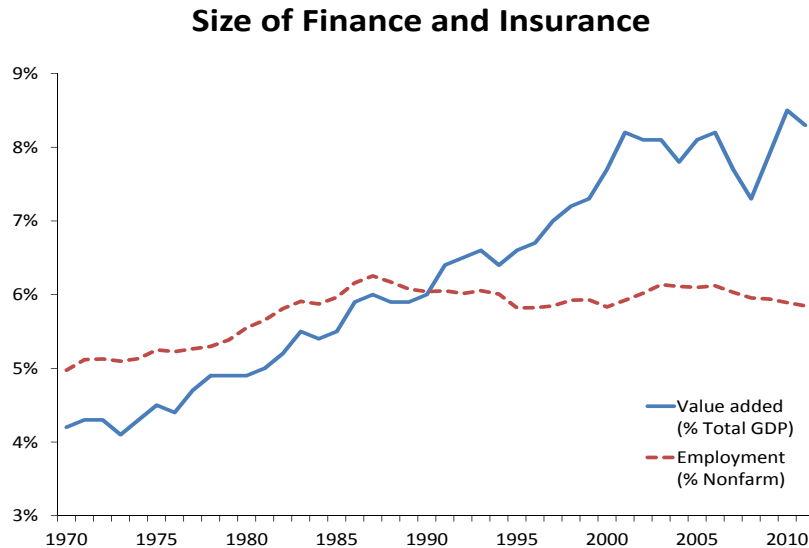


Figure 1: Share of Value Added and Employment

The increase in size was also associated with a sharp increase in compensation. Clementi and Cooley (2009) show that between 1980 and 2007 the average compensation levels in the financial sector increased from parity with other sectors of the economy to 181%. At the same time compensation of managers became more unequal in the financial sector. Figure 2 shows the evolution of the income share of the top 5% of managerial positions in the sector compared to other occupations.

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

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

Income Share of Top 5%

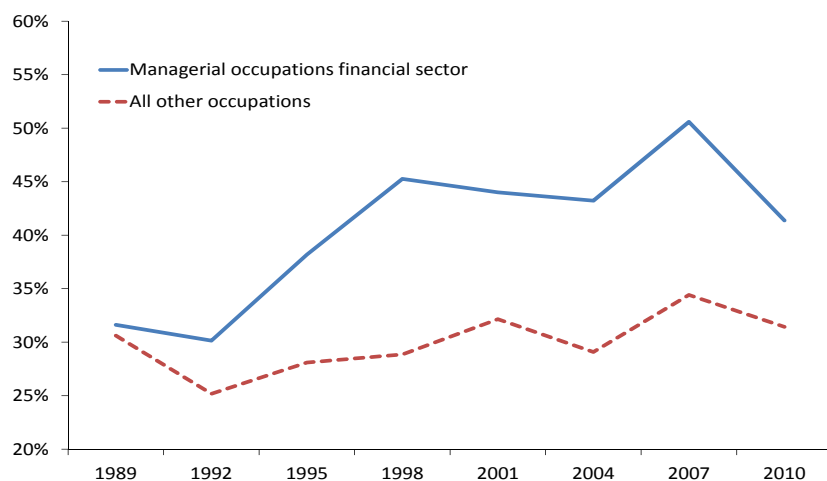


Figure 2: Share of the top 5%

form is characterized by significant agency issues.¹

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. The same evolution occurred in Britain where the closed ownership Merchant Banks virtually disappeared.²

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

¹This is largely consistent with the literature on incomplete contract theory. According to Grossman and Hart (1986) and Hart and Moore (1990), more efficient organizational forms are those where the agents who control the allocation of the investment surplus own a larger share of the assets.

²The fact that member firms were allowed to become public companies does not tell us why they chose to do so. In several cases firms were simply acquired by public companies but in others it was an important strategic decision. Charles Ellis (2008) in his history of Goldman Sachs—the last major firm to go public—suggests that the major motive for financial partnerships to become public was to increase capital for their proprietary trading through an IPO.

significant for the nature of contracts and competition in the financial sector.³

As the structure of financial firms changed, so did the evaluation of them. The market does not seem to value highly the large complex financial institutions. Figure 3 shows the evolution of the ratio of average market value of equity to book value of equity for publicly listed financial and nonfinancial firms since 1970 and shows that, starting in the early 1980's, the market valuation of financial firms has been flat while for nonfinancial firms it has continued to grow. The fact that the market values the financial sector relatively less, compared to the rest of the economy, may be a reflection of compensation practices in firms where managers retain so much of the surplus.⁴

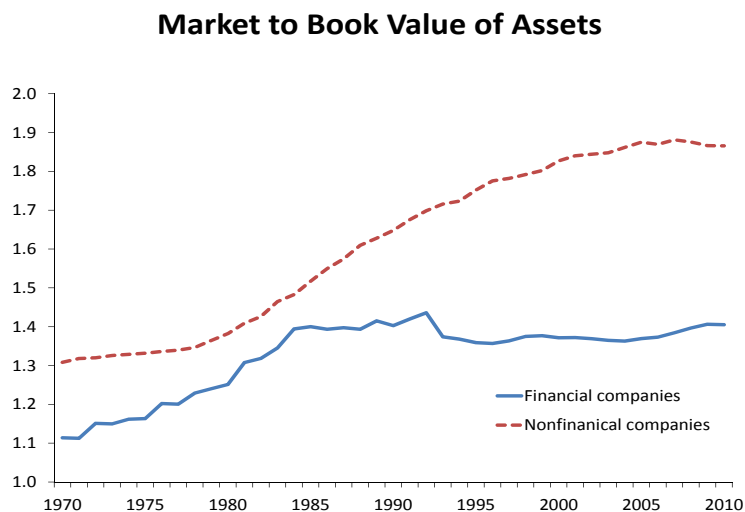


Figure 3: Average Market Value of Equity/Book Values of Equity

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

³Roy 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

⁴Since 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.

accumulated human capital can be transferred outside the firm by the manager, there is a conflict of interest between the investor and the manager. In this environment, the investment desired by the investor may be smaller than the investment desired by the manager because the cost is incurred by the firm while the benefits are shared. This implies that, if the investor cannot control the investment policy either directly or indirectly through a credible compensation scheme, the manager has an incentive to deviate from the optimal policy simply because she does not internalize the full cost of the investment. The goal of the paper is to characterize the investment and compensation policies that result from the (constrained) optimal contract and show how these policies change when the competition for managers increases and the enforcement of contracts weakens.

The basic framework that is often used to study executive compensation is adapted from the principle-agent model of dynamic moral hazard by Spear and Srivastava (1987).⁵ An assumption typically made in this class of models is that the outside option of the agent is exogenous. As argued above, however, an important consequence of the demise of the partnership form is that financial managers are no longer constrained by the limited liquidity of the portion of their wealth that is tied to the firm and it is easier for them to seek outside employment. Since the value of seeking outside employment depends on the market conditions for managers, it becomes important to derive these conditions endogenously in general equilibrium.

A second assumption typically made in principal-agent models is that investors fully commit to the contract. However, as argued above, the clearer separation between investors and managers that followed the transformation of financial partnerships to public companies, could have also reduced the commitment of investors. Therefore, in this paper we relax both assumptions: we endogenize the outside option of managers which will be determined in general equilibrium and we allow for the limited commitment of investors.⁶

To make the outside value of managers endogenous and to study the implications for the whole economy, we embed the micro structure in a general equilibrium model with two sectors—financial and nonfinancial—and two types of workers—skilled and unskilled. Skilled workers can work in both sectors but they innovate only in the financial sector. Unskilled workers have no role in the financial sector. With this general framework we study the consequences of the organizational changes which, as discussed above, had two effects: it increased competition in the financial sector raising the demand for managers and it altered the structure of contractual arrangements between investors and managers

⁵Among the models in this class see, for example, Wang (1997), Quadrini (2004), Clementi and Hopenhayn (2006), Fishman and DeMarzo (2007). Albuquerque and Hopenhayn (2004) can also be considered within this class of models although the frictions are based on limited enforcement rather than information asymmetry.

⁶Although 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 the current paper, instead, uncertainty is central to the analysis. The focus is on risk-taking and how this is affected by changes in the organization of financial firms.

in ways that weakened commitments. These two effects are formalized in the model by a lower cost to create jobs in the financial sector and by a shift to a regime where investors do not commit to the contract (double-sided limited commitment). We then show that these structural changes can generate (i) greater risk-taking; (ii) larger share (and higher relative productivity) of the financial sector; (iii) lower stock market valuation of financial institutions; (iv) greater income inequality within and between sectors.

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

2 The model

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

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

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

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

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

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

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

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

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

If implemented, the new project enhances the human capital of the manager according to $h_{t+1} = h_t + i_{t+1}$. However, for the manager's human capital to grow, it is essential that the new project is implemented. This would happen if the manager continues to work in a financial firm. If the new project is not implemented—for instance, if the manager loses occupation or leaves the financial sector—the human capital of the manager remains h_t . Therefore, if new projects are implemented after their development stage, they become *embedded* human capital. Otherwise they fully depreciate. The importance of this assumption will become clear later.⁷

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

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

The expected lifetime utility of managers takes the form

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

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

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

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

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

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

⁷The assumption that newly developed projects depreciate if not implemented while the pre-existing human capital does not depreciate is not essential for the qualitative properties of the model. It is only made to have linear homogeneity in h_t . This property does not hold if we assume that the whole human capital depreciates (old and new) when the manager moves away from the financial sector.

who chooses to quit at the beginning of period $t + 1$ will be $h_t + i_{t+1}$ only if she receives an offer. Otherwise, the human capital remains h_t .

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

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

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

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

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

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

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

2.1 Optimal contract with one-sided limited commitment

We first characterize the optimal contract when the investor commits but the manager does not (with one-sided limited commitment). In this environment the manager could quit the firm at any point in time and could choose any investment λ_t . The optimal contract can be characterized by solving a planner's problem that maximizes the weighted sum of utilities for the investor and the manager but subject to a set of constraints. These constraints guarantee that the allocation chosen by the planner is enforceable in the sense that both parties choose to participate and the manager has no incentive to take actions other than those prescribed by the contract. We first characterize the key constraints and then we specify the optimization problem.

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

$$E_{t+1} \sum_{n=0}^{\infty} \beta^n \left[u(C_{t+1+n}) - e(\lambda_{t+1+n}) \right] \geq D(h_t, h_{t+1}, \rho), \quad t \geq 0. \quad (3)$$

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

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

$$\hat{\lambda}_t = \arg \max_{\lambda \in [0,1]} \left\{ -e(\lambda) + \beta E_t D(h_t, g(\lambda, \varepsilon_{t+1})h_t, \rho) \right\}. \quad (4)$$

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

$$e_\lambda(\hat{\lambda}_t) \geq \beta E_t D_2(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho) g_\lambda(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \quad (5)$$

which is satisfied with equality if $\hat{\lambda}_t > 0$. We can now see the importance of the assumption that the manager faces the effort dis-utility from innovating. In absence of this, the optimal deviation $\hat{\lambda}_t$ would be 1. Instead, with $e_\lambda(1) = \infty$, the optimal deviation is interior in the interval $[0, 1]$ and is affected by a change in the outside value $D_2(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho)$.

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

$$-e(\lambda_t) + \beta E_t \sum_{n=0}^{\infty} \beta^n \left(u(C_{t+n+1}) - e(\lambda_{t+n+1}) \right) \geq -e(\hat{\lambda}_t) + \beta E_t D(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1})h_t, \rho). \quad (6)$$

Notice that C_t does not appear in (6) because current consumption cannot be contingent on current investment λ_t (since the investment becomes public information in the next period). This limits the ability of the planner to punish the manager by cutting consumption in the current period. The manager can be punished in the next period, once the investment becomes public information. At this stage, however, the manager has always the option to quit, which imposes a lower bound to the possible punishment.

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

write the planner's problem as

$$\begin{aligned} \max_{\{C_t, \lambda_t\}_{t=0}^{\infty}} \quad & E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left(\beta y(\lambda_t) h_t - C_t \right) + \tilde{\mu}_0 \sum_{t=0}^{\infty} \beta^t u(C_t) \right\} \\ \text{s.t.} \quad & (2), (3), (6). \end{aligned} \quad (7)$$

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

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

$$\begin{aligned} W(h, \tilde{\mu}) = \min_{\tilde{\chi}, \tilde{\gamma}(\varepsilon')} \max_{C, \lambda} \quad & \left\{ \beta y(\lambda) h - C + \tilde{\mu} \left(u(C) - e(\lambda) \right) - \tilde{\chi} \left(e(\lambda) - e(\hat{\lambda}) \right) + \right. \\ & \left. \beta E \left[W(h', \tilde{\mu}') - \tilde{\chi} D \left(h, g(\hat{\lambda}, \varepsilon') h, \rho \right) - \tilde{\gamma}(\varepsilon') D(h, h', \rho) \right] \right\} \\ \text{s.t.} \quad & h' = g(\lambda, \varepsilon') h, \quad \tilde{\mu}' = \tilde{\mu} + \tilde{\chi} + \tilde{\gamma}(\varepsilon'), \end{aligned} \quad (8)$$

where $\tilde{\gamma}(\varepsilon')$ is the Lagrange multiplier for the enforcement constraint (3) and $\tilde{\chi}$ is the Lagrange multiplier for the incentive-compatibility constraint (6).

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

Differentiating problem (8) by C we obtain the optimality condition

$$C_t = u_c^{-1} \left(\frac{1}{\tilde{\mu}_t} \right), \quad (9)$$

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

It is useful to consider the *normalized* manager's weight, or manager's share of the surplus, $\mu_t = \tilde{\mu}_t/h_t$. Since, income is proportional to h_t (recall (1)), this inverse relation between the manager's share and her human capital plays a key role in our analysis. In particular, $Eh_{t+1} = (1 + p\lambda_t)h_t$. Therefore, as long as the contract prescribes $\lambda_t > 0$, then $Eh_{t+1} > h_t$, that is, human capital increases on average over time. If contracts were perfectly enforceable, $\tilde{\chi}_t$ and $\tilde{\gamma}_t(\varepsilon_{t+1})$ would be both equal to zero and $\tilde{\mu}_{t+1} = \tilde{\mu}_t$. This implies that the normalized share of the surplus μ_t converges to zero. However, with one-sided limited commitment, the enforcement and incentive-compatibility constraints set a lower bound on μ_t . In particular, $\tilde{\mu}_{t+1}$ increases when either the enforcement constraint binds ($\tilde{\gamma}_t(\varepsilon_{t+1}) > 0$) or the incentive-compatibility constraint binds ($\tilde{\chi}_t > 0$). Thus, with one-sided limited commitment the manager has a minimum share guaranteed. The

consequence of this is that the contract does not provide full insurance since the manager's consumption increases stochastically.

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

$$e_{\lambda}(\lambda_t) \left(\frac{\tilde{\mu}_t + \tilde{\chi}_t}{h_t} \right) - \beta y_{\lambda}(\lambda_t) \geq \beta E_t \left[W_1 \left((1 + \lambda_t \varepsilon_{t+1}) h_t, \tilde{\mu}_{t+1} \right) - \tilde{\gamma}_t(\varepsilon_{t+1}) D_2 \left(h_t, g(1 + \lambda_t \varepsilon_{t+1}) h_t, \rho \right) \right] \varepsilon_{t+1}, \quad (10)$$

which is satisfied with equality if $\lambda_t > 0$.

The left-hand side of (10) is the marginal cost of investment per unit of human capital. This is increasing in λ_t , μ_t and $\chi_t = \tilde{\chi}_t/h_t$. The right-hand-side is the expected benefit from investing. With full commitment, λ_t is increasing because $\chi_t = 0$, $\tilde{\mu}_t = \tilde{\mu}_0$ and μ_t converges to zero. With limited commitment, however, the limited enforcement and/or the incentive compatibility constraints could be binding, raising the marginal cost. Looking now at the marginal benefit, we can see that the increase in W_1 induced by the increase in human capital is counterbalanced by the increase in D_2 . Thus, the effect of binding constraints is a lower λ_t (lower risk), since the marginal cost incurred by the manager is higher (when $\tilde{\gamma}_t(\varepsilon_{t+1}) + \tilde{\chi}_t > 0$) and the marginal benefit is curtailed by the fact that the outside value is higher (when $\gamma_t(\varepsilon_{t+1}) > 0$). More importantly, since $D_{2,3} > 0$, an increase in the degree of competition as captured by the parameter ρ , increases the impact of $\gamma_t(\varepsilon_{t+1}) > 0$. Formally,

Proposition 1 *Suppose that the optimal investment is interior, that is $\lambda_t^* \in (0, 1)$. More competition for managers (higher ρ) affects investment by increasing the weight given to the manager when the limited enforcement and incentive compatibility constraints are binding. It affects investment directly only when the enforcement constraint is binding, in which case it lowers λ_t^* .*

2.2 Optimal contract with double-sided limited commitment

The law of motion $\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1})$ captures the investor's commitment to fulfill promises made to the manager. With double-sided limited commitment the investor does not commit to fulfill his promises and renegotiates whenever the value of the contract for the manager exceeds the outside value. This implies that the value of $\tilde{\mu}_t$ chosen in the previous period becomes irrelevant for the new $\tilde{\mu}_{t+1}$ chosen in the current period. Under these conditions, the manager has the incentive to choose the investment that maximizes the outside value as defined in (4), that is, $\lambda_t = \hat{\lambda}_t$. Thus, the incentive-compatibility

constraint and the multiplier $\tilde{\chi}_t$ become irrelevant, and the optimal contract solves

$$\begin{aligned}
W(h, \tilde{\mu}) = \min_{\tilde{\gamma}(\varepsilon')} \max_C & \left\{ \beta y(\hat{\lambda})h - C + \tilde{\mu} \left(u(C) - e(\hat{\lambda}) \right) + \right. \\
& \left. \beta E \left[W \left(g(\hat{\lambda}, \varepsilon')h, \tilde{\mu}' \right) - \tilde{\gamma}(\varepsilon') D \left(h, g(\hat{\lambda}, \varepsilon')h, \rho \right) \right] \right\} \quad (11) \\
\text{s.t.} \quad & \tilde{\mu}' = \tilde{\gamma}(\varepsilon').
\end{aligned}$$

The contract with double-sided limited commitment simply prescribes a consumption plan which is determined by (9) with $\tilde{\mu}' = \tilde{\gamma}(\varepsilon')$, and the investment is $\hat{\lambda}$, which is the solution to (5). Since $D_{2,3} > 0$, an increase in competition captured by the parameter ρ increases the right-hand-side of (5), that is, the the marginal benefit of investing for the manager. This is stated formally in the next proposition.

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

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

2.3 The normalized contract

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

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

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

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

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

The value of the contract for the investor can be expressed recursively as $V_t = \beta y(\lambda_t)h_t - C_t + \beta E_t V_{t+1}$ and normalized to

$$v_t = \beta y(\lambda_t) - c_t + \beta E_t g(\lambda_t, \varepsilon_{t+1})v_{t+1}, \quad (12)$$

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

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

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

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

$$q_t = \ln(c_t) + \alpha \ln(1 - \lambda_t) + \beta E_t \left[\mathcal{B} \ln\left(g(\lambda_t, \varepsilon_{t+1})\right) + q_{t+1} \right]. \quad (13)$$

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

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

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 (6) can be written as

$$q_{t+1} \geq (1 - \rho) \underline{q} + \rho \bar{q} - (1 - \rho) \mathcal{B} \ln\left(g(\lambda_t, \varepsilon_{t+1})\right). \quad (14)$$

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

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

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

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

$$\begin{aligned} \alpha \ln(1 - \lambda_t) + \beta E_t \left[q_{t+1} + \mathcal{B} \ln \left(g(\lambda_t, \varepsilon_{t+1}) \right) \right] &\geq \\ \alpha \ln(1 - \hat{\lambda}_t) + \beta E_t \left[(1 - \rho) \underline{q} + \rho \bar{q} + \rho \mathcal{B} \ln \left(g(\hat{\lambda}_t, \varepsilon_{t+1}) \right) \right]. \end{aligned} \quad (15)$$

We can now provide a more explicit characterization of the manager's optimal investment deviation $\hat{\lambda}_t$. This maximizes the expected value of quitting net of the effort cost, that is, the right-hand-side of (15). Using $g(\lambda, \varepsilon) = 1 + \lambda \varepsilon$ and remembering that $\varepsilon \in \{0, 1\}$ with probabilities $1 - p$ and p , the optimality condition (5) can be written as

$$\frac{\alpha}{1 - \hat{\lambda}_t} \geq \frac{\rho \beta \mathcal{B} p}{1 + \hat{\lambda}_t}, \quad (16)$$

which is satisfied with equality if $\hat{\lambda}_t > 0$. As implied by Proposition 2, we can now see more explicitly that $\hat{\lambda}$ is increasing in the probability ρ .

The original contractual problem (7) with one-sided limited commitment can be reformulated in normalized form using the 'promised utility' approach. This maximizes the normalized investor's value subject to the normalized promise-keeping, limited enforcement and incentive-compatibility constraints, that is,

$$v(q) = \max_{\lambda, c, q(\varepsilon')} \left\{ \beta y(\lambda) - c + \beta E g(\lambda, \varepsilon') v(q(\varepsilon')) \right\} \quad (17)$$

subject to (13), (14), (15).

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

Policies $\varphi^\lambda(q)$, $\varphi^c(q)$, and $\varphi^q(q, \varepsilon)$ satisfy the first order conditions (Appendix A):

$$c = \mu, \quad (18)$$

$$\frac{\alpha(\mu + \chi)}{1 - \lambda} - \beta y_\lambda(\lambda) = \beta p \left[v(q(\bar{\varepsilon})) + \frac{\mathcal{B}[\mu + \chi + (1 - \rho)\gamma(\bar{\varepsilon})]}{1 + \lambda} \right], \quad (19)$$

$$\mu(\varepsilon) = \frac{\mu + \chi + \gamma(\varepsilon)}{1 + \lambda \varepsilon}. \quad (20)$$

The variables μ , $\gamma(\varepsilon)$ and χ are the Lagrange multipliers for constraints (13)-(15). The envelope condition $v'(q) = -\mu$ shows the equivalence between the normalized problem (17) and the original problem (7).⁸

For the case with double-sided limited commitment, we can reformulate problem (11) in normalized form in a similar fashion. Using the ‘promised utility’ approach, the partnership contract with double-sided limited commitment can be written as

$$v(q) = \max_{c, q(\varepsilon)} \left\{ \beta y(\hat{\lambda}) - c + \beta E g(\hat{\lambda}, \varepsilon) v(q(\varepsilon)) \right\} \quad (21)$$

subject to

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

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

where $\hat{\lambda}$ is determined by condition (16).

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

$$q_{min} = \ln(c(q_{min})) + \alpha \ln(1 - \hat{\lambda}) + \beta E_t \left[(1 - \rho)q + \rho \bar{q} + \rho \mathcal{B} \ln(g(\hat{\lambda}, \varepsilon)) \right]. \quad (22)$$

Problem (21) can be seen as a special case of problem (17) where we have replaced the incentive-compatibility constraint (15) with $\lambda = \hat{\lambda}$. Furthermore, we have imposed that the enforcement constraint (14) is always satisfied with equality. This is because any promise that exceeds the outside value of the manager will be renegotiated ex-post. Notice that in this problem the decision variables, c and $q(\varepsilon)$, are fully determined by the promise-keeping and incentive-compatibility constraints. Therefore, the problem can be solved without performing any optimization, besides solving for $\hat{\lambda}$.

2.4 Contract properties

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

As we have seen, the solution to the contractual problem (17) with one-sided limited commitment provides the optimal policies for investment, $\lambda = \varphi^\lambda(q)$, manager’s

⁸Appendix B shows another way of writing the optimization problem recursively, starting directly from the original problem (7). From the recursive problem we obtain the same first-order conditions (18) and (19) while condition (20) is simply the law of motion of the co-state variable in (7).

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 h , we can construct the whole sequence of human capital 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 (18)-(20). This is also the case for the solution to problem (21) in the environment with double-sided limited commitment.

Figure 4 plots the values of next period normalized continuation utilities, $q(\varepsilon) = \varphi^q(q, \varepsilon)$, and investment λ_t , as functions of current normalized utility, q . The initial ‘normalized’ value of the contract for the manager, \bar{q} , is shown by the vertical line. How this value is determined will be described later when we extend the model to a general equilibrium. The left panels are for the environment with one-sided limited commitment and the right panels are for the environment with double-sided limited commitment.

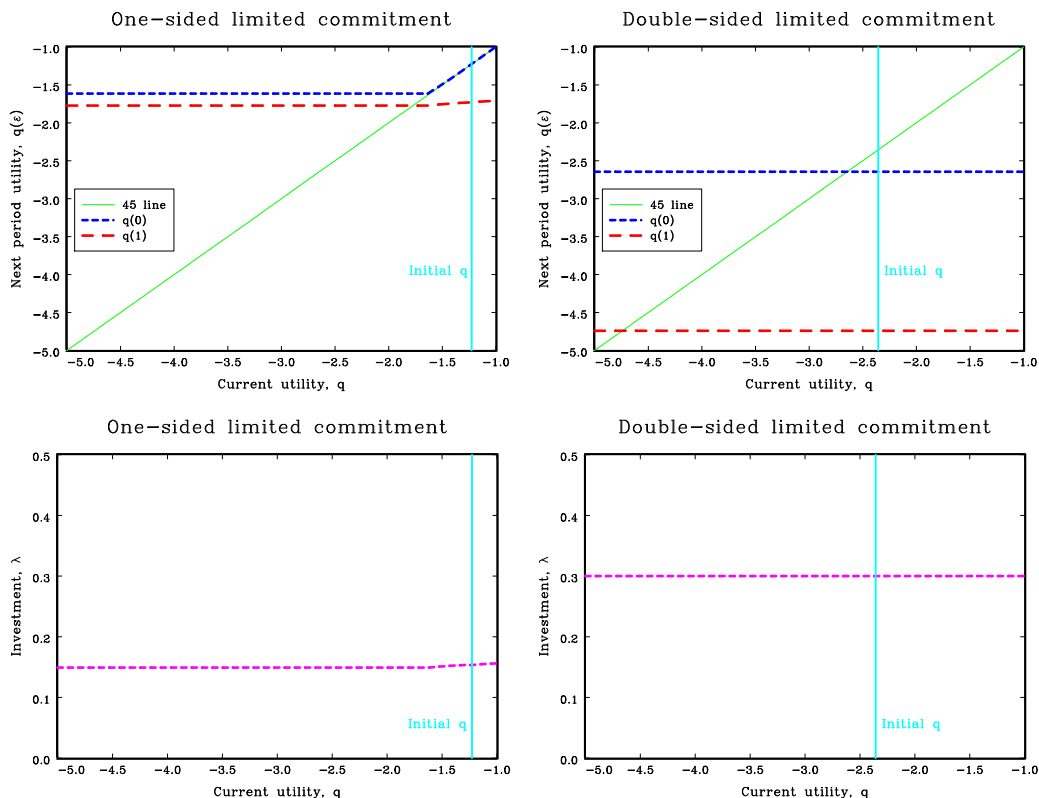


Figure 4: Continuation utilities and investment with one-sided and double-sided limited commitment.

The dynamics of promised utilities. 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. Therefore, q declines on average until it reaches a lower bound.

This dynamics 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. When the enforcement constraint is not binding, it becomes optimal to keep the 'normalized' utility constant when the investment fails ($\varepsilon = 0$) and to decrease it when the investment succeeds ($\varepsilon = 1$). To show this, let's consider the non-normalized manager's utility which satisfies

$$Q(h) = \ln(C) + \alpha \ln(1 - \lambda) + \beta \left[pQ((1 + \lambda)h) + (1 - p)Q(h) \right].$$

When the enforcement constraint is not binding, $C = ch$ is constant and λ increases with h (for higher values of human capital it becomes efficient to exert more effort to innovate). Therefore, $Q(h) > Q((1 + \lambda)h)$. Since the manager's value is normalized as $q = Q(h) - \mathcal{B} \ln(h)$, it follows that

$$\begin{aligned} q(1) &= Q((1 + \lambda)h) - \mathcal{B} \ln((1 + \lambda)h) \\ &< Q(h) - \mathcal{B} \ln(h) = q(0). \end{aligned}$$

However, as q declines, the limited enforcement constraint becomes binding and, eventually, the normalized utility reaches the lower bound q_{min} defined in (22). At this point consumption grows at the same rate as h , while the normalized utility q remains constant at q_{min} .

In the environment with double-sided limited commitment (see right panel of Figure 4), 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 value and fluctuates between two values. The fact that the initial q is bigger than future values implies that in the first period the manager receives a higher payment (consumption) relatively 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 to a higher promised utility. Since this is costly for the investor, the optimal λ is lower for lower values of q (although quantitatively the dependence is small) until it reaches the lower bound q_{min} .

We now turn to the environment with double-sided limited commitment. In this case λ is independent of q since the manager always choose $\lambda = \hat{\lambda}$. Given the limited commitment of the investor, the manager knows that her value is always equal to the

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 in the double-sided limited commitment is bigger than in the environment with one-sided limited commitment. However, this property cannot be generalized to any set of parameters because there are two contrasting effects. On the one-hand, with double-side limited commitment, the manager does not take into account the loss of production. This leads to a higher λ . On the other, when the manager chooses λ , she maximizes the outside value, which is the value of finding occupation in another firm. But the investment has value only if the manager finds a new occupation, which happens with probability $\rho < 1$. Instead, the investment made within the current firm has value with probability 1. This leads the manager to choose a lower value of λ . Therefore, to have that the investment in the double-sided limited commitment is bigger than the investment with one-sided limited commitment, we need that the marginal production loss from innovation (the derivative of $y(\lambda)$) and the probability of finding outside occupation (the probability ρ) are sufficiently large.

3 General model

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

There are two sectors in the model—financial and nonfinancial—and three types of agents—a unit mass of *investors*, a unit mass of *skilled workers*, and a mass $N > 1$ of *unskilled workers*. Unskilled workers are only employed in the nonfinancial sector while skilled workers can be employed in either sectors.⁹

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, skilled and unskilled, have the same utility $\ln(c_t) + \alpha \ln(1 - \lambda_t)$. However, only managerial occupations in the financial sector requires effort λ_t and, therefore, the utility of unskilled and skilled workers employed in the nonfinancial sector reduces to $\ln(c_t)$.

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 skilled workers are endowed with initial human capital h_0 while the human capital of unskilled workers is normalized to 1. The motivation for adding this particular demographic structure is to prevent the distribution of h_t to become degenerate. The assumption of a constant h_0 together with the finite lives of skilled workers guarantee that the distribution of h_t across financial managers converges to an invariant distribution and the model is stationary in level. Taking into

⁹An 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. However, in this paper we prefer to focus on the financial sector because the changes described in the introduction are more evident.

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 any modification.

In every period, a fraction ψ of new born skilled workers have the ability to become managers in the financial sector. We denote by S the total mass of skilled workers employed in the nonfinancial sector (with and without ability to become financial managers) and $1 - S$ is the mass of skilled workers employed in the financial sector. The assumption that only a fraction ψ of skilled 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 $F(N, H) = K^\nu H^{1-\nu}$, where N is the number of unskilled workers and H is the aggregate efficiency-units of labor supplied by skilled workers who are employed in the nonfinancial sector. This results from the aggregation of human capital of all skilled workers employed in the nonfinancial sector. As we will see, in equilibrium, the human capital of skilled workers 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 rates earned in the nonfinancial sector by unskilled and skilled workers are equal to marginal productivities, that is,

$$w^N = \nu \left(\frac{N}{H} \right)^{\nu-1}, \quad w^S = (1 - \nu) \left(\frac{N}{H} \right)^\nu. \quad (23)$$

While the nonfinancial sector is competitive, the hiring process in the financial sector is characterized by matching frictions. More specifically, the skilled workers who have 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 skilled worker currently employed in the nonfinancial sector with human capital h and with the ability to become a financial manager is

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

The skilled worker consumes income $w^S h$ in the current period. In the next period, with probability ρ_{t+1} she finds an occupation in the financial sector. In this case the lifetime utility is $\bar{Q}_{t+1}(h)$. 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 $\bar{Q}_{t+1}(h)$ is the value of a new contract for the financial manager. Therefore, the probability ρ_{t+1} and the outside values $\underline{Q}_{t+1}(h)$ and $\bar{Q}_{t+1}(h)$ are now endogenous and determined in the general equilibrium.

3.1 Matching and general equilibrium

In the financial sector, investors post vacancies for skilled workers with managerial abilities. A vacancy specifies the level of human capital h and the value of the contract for

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

Investors can freely post vacancies in the financial sector and the following free-entry condition will be satisfied in equilibrium:

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

We can now take advantage of the properties of the optimal contract characterized in the previous section where we have shown that the value of the contract for the investor is linear in h , that is, $V_t(h, \bar{Q}_t) = v_t(\bar{q}_t)h$. The variable \bar{q}_t is the normalized value of the contract for a newly hired manager. Therefore, using the normalization, it is enough to determine \bar{q}_t to define a menu of posted contracts for all possible levels of human capital h . More precisely, once \bar{q}_t is decided, the investor offers $\bar{Q}_t = \bar{q}_t + \mathcal{B} \ln(h)$ to the manager 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

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

Appendix C discusses the equilibrium conditions in more detail and shows that in equilibrium the manager receives a fraction $1 - \eta$ of the matching surplus. This is the standard efficiency property of directed search models. As it 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 for skilled workers in the nonfinancial sector, equation (24). This can be rewritten as

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

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

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 skilled workers in the nonfinancial sector, \underline{q} , skilled workers newly hired in the financial sector, \bar{q} , and initial normalized value for investors, \bar{v} ;
3. Skilled workers employed in the nonfinancial sector, S , and those with managerial ability, U_t . Posted vacancies X , filling probability, ϕ , and finding probability, ρ ;
4. Wage rates in the nonfinancial sector, w^N and w^S .
5. Distribution of skilled workers employed in the financial sector $\mathcal{M}(h, q)$;
6. Law of motion for the distribution of financial managers, $\mathcal{M}_{t+1} = \Phi(\mathcal{M}_t)$;

Such that

1. The policy rules $\varphi^\lambda(q)$, $\varphi^c(q)$, $\varphi^q(q, \varepsilon)$ solve the optimal contract;
2. The normalized utilities \underline{q} and \bar{q} and investor value \bar{v} solve (25), (26) and (34);
3. Filling and finding probabilities satisfy $\phi = m(X, U)/X$ and $\rho = m(X, U)/U$.
4. The wage rates in the nonfinancial sector are marginal productivities as in (23).
5. The law of motion $\Phi(\mathcal{M})$ is consistent with contract policies $\varphi^\lambda(q)$ and $\varphi^q(q, \varepsilon)$.
6. The distribution of managers is constant, that is, $\mathcal{M} = \Phi(\mathcal{M})$.

It will be convenient for the later analysis to state formally the property for which increasing competition for managers redistributes rents in their favour. The proof is provided in Appendix C.

Lemma 3 *The contract value \bar{q} offered to the manager is increasing in ρ .*

3.2 Inequality

The general model features three types of workers: 1) unskilled workers employed in the nonfinancial sector; 2) skilled workers employed in the nonfinancial sector (some of which with the ability to become a manager in the financial sector); 3) skilled workers employed in the financial sector. Therefore, we can study inequality between skilled and unskilled workers; between skilled workers employed in the financial and nonfinancial sectors; across skilled workers employed in the financial sector. Here we focus on the distribution of income across skilled workers employed in 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 skilled 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 of skilled 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 financial occupation,

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

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

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

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

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

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

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

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

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

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

Therefore, 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 $\hat{\lambda}$, instead, can be explained as follows. If $\hat{\lambda} = 0$, the human capital of all managers will be equal to h_0 and the inequality index is zero. As $\hat{\lambda}$ becomes positive,

inequality increases for two reasons. First, since the growth rate $g(\hat{\lambda}, \varepsilon)$ is stochastic, human capital will differ *within* the same tenure cohort of managers (managers with the same employment tenure). Second, since each cohort experiences growth, the average human capital differs *between* cohorts of managers. More importantly, the cross sectional dispersion in human capital induced by these two mechanisms (the numerator of the inequality index) dominates the increase in average human capital (the denominator of the inequality index). Thus, inequality increases in $\hat{\lambda}$.

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

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

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

4 The impact of organizational changes

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

1. ***Increased competition for managers:*** The separation between investors and managers enlarged the base of potential investors who could fund new investment projects. In the context of our model this is captured, parsimoniously, by a reduction in the vacancy cost τ . A lower value of τ generates more entry 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 and contracts have limited enforceability also for investors. We start exploring the consequences of the higher 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. *Higher share and relative productivity of the financial sector.*
3. *Lower stock market valuation of financial institutions.*
4. *Greater income inequality within and between sectors (financial and nonfinancial).*

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 the manager to exert more effort to increase the outside value.

The second property, that is, the increase in the size of the financial sector derives in part from higher employment and in part from higher investment. The increase in the share of employment, however, would arise even if there were no contractual frictions. The higher investment, instead, is a direct consequence of the 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 this figure, the share of value added of the financial sector has increased much more than the share of employment.

The third property 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 where there is a division between investors and managers. However, the effect is likely to be stronger when there is limited commitment also for investors. We will show this in the numerical simulation.

Finally, the fourth property 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 in the financial sector. At the same time, since more skilled workers will be employed in the financial sector, fewer skilled workers will be employed in the nonfinancial sector. Because of the complementarity between skilled and unskilled workers in the nonfinancial sector, the wage of skilled workers increases while the wage of unskilled workers declines. Thus, the model generates greater inequality also between sectors and between skills.

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. For this change, however, we cannot derive analytical results. Therefore, we conduct a quantitative analysis.

4.1 Quantitative analysis

We calibrate the model annually and use the 2000s to measure the calibration targets. 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 which we assume to be quadratic, that is, $\xi(\lambda) = 1 - \lambda^2$. Therefore, if the manager devotes all of her time producing ($\lambda = 0$), each unit of human capital produces one unit of output. If instead the manager allocates all of her time innovating ($\lambda = 1$), production is zero.

Given the specification of preferences and technology, there are 11 parameters to calibrate as listed in the top section of Table 1. Given the difficulties to calibrate the parameter of the matching function η , it is customary to set it to $\eta = 0.5$. We follow the same approach here even if in our model jobs are created through matching only in the financial sector. We are then left with 10 parameters which we calibrate using the 10 moments listed in the bottom section of Table 1.

Table 1: Parameters and calibration moments.

<i>Parameters</i>		
$\hat{\beta}$	Discount factor	0.962
ω	Death probability	0.025
N	Number unskilled workers	1.000
ψ	Fraction of skilled workers searching for financial jobs	0.168
p	Probability of successful innovation	0.035
α	Utility parameter for dis-utility innovation effort	0.139
ν	Production parameter in the nonfinancial sector	0.704
h_0	Human capital of newborn skilled workers	0.643
τ	Cost of posting a vacancy in the financial sector	0.320
A	Matching productivity	1.000
η	Matching share parameter (pre-set)	0.500
<i>Calibration moments</i>		
	Interest rate	0.04
	Life expectancy of workers	40.00
	Fraction of skilled workers	0.25
	Skill premium in the nonfinancial sector	0.50
	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 7 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 25%

of skilled workers is the approximate number for the fraction of college graduates in the United States. 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. In absence of more specific information, we assign these numbers arbitrarily. However, a sensitivity analysis will be conducted with respect to these three parameters. Appendix F provides a detailed description of how the 10 moments are mapped into the parameters.

Results. The goal of this section is to study the quantitative impact of higher competition and lower enforcement. The impact of higher competition is studied by looking at the equilibrium consequences of reducing the vacancy cost τ . The impact of lower enforcement is studied 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 τ as characterizing the financial sector in the pre-1980s period. The environment with double-sided limited commitment and lower τ , instead, as representative of the most recent years.

Since the 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. We will then show the results for alternative values of τ .

Figure 5 plots the steady state policy $\lambda = \varphi^\lambda(q)$ in the environments with one-sided and double-sided limited commitment, and for the two values of τ . In the environments with one-sided limited commitment, more competition (lower τ) reduces slightly the investment λ , which is consistent with Proposition 1. 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. Quantitatively, however, the change in λ is very small.

In contrast, when neither managers nor investors can commitment, more competition induces higher 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 values is by raising their outside value. This is achieved by choosing higher λ . With a lower τ , however, 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 .

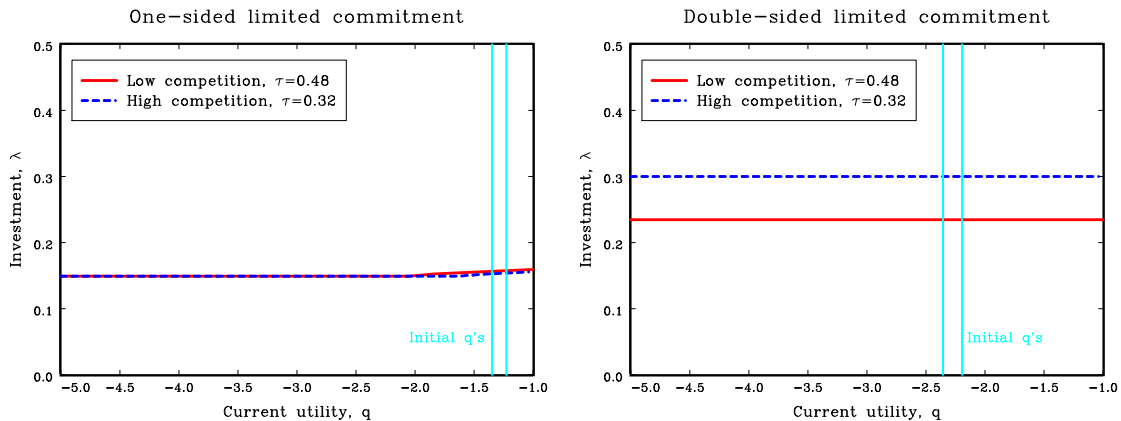


Figure 5: Steady state investment policies for different τ in the environments with one-sided and double-sided limited commitment.

So far we have shown that the organizational changes that took place in the financial sector induced higher risk-taking. We now show that they also generated other changes observed in the US economy. As shown in Table 2, the shift to an environment with double-sided limited commitment and lower τ is associated with only a small increase in the share of employment in the financial sector but a significant increase in the share of output generated by the financial sector. The increase in employment share is only 0.1% but the increase in output share is 2.5%.

Another 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.¹⁰ Table 2 also shows that the *interim* investor's value and the probability of filling a vacancy are both lower. This follows directly from the free entry condition $\phi(\bar{q}) \cdot v(\bar{q}) = \tau$ after the reduction in the vacancy cost τ . Furthermore, it also shows how the decrease in \bar{v} is larger in the environment with double-sided limited commitment.

These two 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 in the financial sector, which is consistent with the evidence provided by Figure 2. The model generates also some increase in inequality in the nonfinancial sector between skilled and non-skilled workers. However, the increase is not very large because the movements of skilled workers to the financial sector is small. This, in turn, follows for the fact that the calibrated value of the parameter ψ is only 0.168. As a result, there are only a small number of skilled workers with managerial ability who can move to the financial sector. Therefore, even if the probability ρ increases significantly (from 0.4 to 0.5), the number of skilled workers leaving the nonfinancial sector is small.

¹⁰This would be the case if we explicitly introduce capital and assume that there is complementarity between human and physical capital.

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

	<i>One-sided limited commitment</i>	<i>Double-sided limited commitment</i>
Low competition ($\tau = 0.480$)		
Average value of λ	0.154	0.235
Offer probability, ρ	0.396	0.434
Filling probability, ϕ	0.632	0.576
Share of employment financial sector	0.039	0.040
Share of output financial sector	0.065	0.072
Earnings unskilled workers	0.424	0.424
Earnings skilled workers nonfinancial sector	0.635	0.636
Earnings skilled workers financial sector	0.698	0.775
Initial investor value \bar{v}	0.760	0.834
Average investor value $Ev(q)$	1.054	1.249
Within inequality fin sector	0.060	0.311
Between inequality fin sector	0.076	0.280
Coefficient of variation	0.369	0.769
High competition ($\tau = 0.320$)		
Average value of λ	0.151	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 unskilled workers	0.424	0.424
Earnings skilled workers nonfinancial sector	0.636	0.636
Earnings skilled workers financial sector	0.706	0.871
Initial investor value \bar{v}	0.636	0.640
Average investor value $Ev(q)$	0.803	0.948
Within inequality fin sector	0.058	3.110
Between inequality fin sector	0.073	0.890
Coefficient of variation	0.362	2.000

5 Conclusion

The financial crisis of 2007-2009 has brought attention to the growth in size and importance of the financial sector over the past few decades as well as the increase in risk taking. 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. These changes could have also played an important role in another widely documented trend occurred during the same period—the increase in income inequality.

The fact that inequality has increased over time, especially in anglo saxon countries,

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

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

Appendix

A First order conditions for Problem (17)

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

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

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

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

Substituting the envelope condition $v_q(q) = -\mu$ and using the functional forms of $y(\lambda)$ and $g(\lambda, \varepsilon)$ we obtain equations (19)-(20).

B Alternative formulation of the normalized problem

The contractual Problem (8) for one-sided limited commitment can be normalized as

$$\begin{aligned}
w(\mu) &= \min_{\chi, \gamma(\varepsilon')} \max_{c, \lambda} \left\{ \beta y(\lambda) - c + \mu \left(\ln(c) + \alpha \ln(1 - \lambda) \right) - \chi \alpha \ln \left(\frac{1 - \lambda}{1 - \hat{\lambda}} \right) + \right. \\
&\quad \left. \beta E \left[\mu' \mathcal{B} \ln(g(\lambda, \varepsilon')) + g(\lambda, \varepsilon') w(\mu') - \chi d(\hat{\lambda}, \varepsilon', \rho) - \gamma(\varepsilon') d(\lambda, \varepsilon', \rho) \right] \right\} \\
&\quad \text{s.t.} \quad \mu' = \frac{\mu + \chi + \gamma(\varepsilon')}{g(\lambda, \varepsilon')},
\end{aligned} \tag{29}$$

where $d(\lambda, \varepsilon', \rho) = (1 - \rho)q + \rho\bar{q} + \rho\mathcal{B}\ln(g(\lambda, \varepsilon))$ and $w(\mu) = v^P(\mu) + \mu q^P(\mu)$, with v^P and q^P being the normalized values (as a function of the Pareto weight μ) of the investor and the manager, respectively. That is, $q^P(\mu) = Q(h, \tilde{\mu}) - \mathcal{B}\ln(h)$ and $w(\mu) = W(h, \tilde{\mu})/h - \mu\mathcal{B}\ln(h)$. It is easy to verify that the first order conditions for this problem are given by equations (18)-(20). The contractual problem with double-sided limited commitment can be seen as a special case of (29) with $\lambda = \hat{\lambda}$ and $\chi = 0$, resulting in the law of motion $\mu' = \hat{\gamma}(\varepsilon')/g(\hat{\lambda}, \varepsilon')$.

C The posted contract

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

$$\rho_t \left(h, \bar{Q}_t(h) \right) \cdot \left[\bar{Q}_t(h) - \underline{Q}_t(h) \right] = \rho_t \left(h, \bar{Q}_t^*(h) \right) \cdot \left[\bar{Q}_t^*(h) - \underline{Q}_t(h) \right], \quad (30)$$

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

$$\max_{\bar{Q}_t(h)} \left\{ \phi_t \left(h, \bar{Q}_t(h) \right) \cdot V \left(h, \bar{Q}_t(h) \right) \right\} \quad (31)$$

subject to (30),

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 $V_t(h, \bar{Q}_t(h))$.

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

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

$$\phi_t \left(h, \bar{Q}_t(h) \right) \cdot V_t \left(h, \bar{Q}_t(h) \right) = \tau h. \quad (32)$$

We can take advantage of the of the linear property of the model and normalize the above equations. We have shown that the value of a contract for the investor is linear in h ,

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

$$\phi_t(\bar{q}_t) \cdot v_t(\bar{q}_t) = \tau. \quad (33)$$

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)¹¹.

The investor's problem (31) can be rewritten as

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

subject to

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

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

$$1 - \eta = \frac{-v'_t(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}{v_t(\bar{q}_t) - v'_t(\bar{q}_t)(\bar{q}_t - \underline{q}_t)}. \quad (34)$$

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. $v_\epsilon(\bar{q}) \equiv -\frac{v'(\bar{q})\bar{q}}{v(\bar{q})}$.

Lemma 3A Assume that $v_\epsilon(\bar{q})$ is non-decreasing, then the contract value \bar{q} offered to the manager in a stationary equilibrium is increasing in ρ .

First we notice that with our log-linear specification the assumption is satisfied. Second, we rewrite the optimality condition (34) as

$$\frac{1 - \eta}{\eta} = v_\epsilon(\bar{q}) \frac{\bar{q} - \underline{q}}{\bar{q}}. \quad (35)$$

In a stationary equilibrium

¹¹In 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 .

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

therefore

$$\frac{\bar{q} - \underline{q}}{\bar{q}} = (1 - \beta(1 - \rho))^{-1} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right],$$

and we can rewrite (35) as

$$\frac{1 - \eta}{\eta} = v_\epsilon(\bar{q})(1 - \beta(1 - \rho))^{-1} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right].$$

Taking derivatives with respect to ρ and not considering the (second order) effect on ω :

$$\begin{aligned}
0 &= v'_\epsilon(\bar{q})(1 - \beta(1 - \rho))^{-1} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right] \bar{q}'(\rho) - \beta v_\epsilon(\bar{q})(1 - \beta(1 - \rho))^{-2} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right] \\
&\quad + v_\epsilon(\bar{q})(1 - \beta(1 - \rho))^{-1} \frac{\ln(\omega)}{\bar{q}^2} \bar{q}'(\rho)
\end{aligned}$$

i.e.

$$\beta(1 - \beta(1 - \rho))^{-1} = \left\{ \frac{v'_\epsilon(\bar{q})}{v_\epsilon(\bar{q})} + \frac{\ln(\omega)}{\bar{q}^2} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right]^{-1} \right\} \bar{q}'(\rho).$$

Since $\ln(\omega) < \bar{q} < 0$, $\left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right] < 0$ and the term within brackets is positive, whenever $v'_\epsilon(\bar{q}) \geq 0$, and therefore $\bar{q}'(\rho) > 0$. If the (second order) effect on ω is taken into account, the previous equation is modified to

$$\beta(1 - \beta(1 - \rho))^{-1} + \frac{\omega'(q)}{\omega\bar{q}} = \left\{ \frac{v'_\epsilon(\bar{q})}{v_\epsilon(\bar{q})} + \frac{\ln(\omega)}{\bar{q}^2} \left[(1 - \beta) - \frac{\ln(\omega)}{\bar{q}} \right]^{-1} \right\} \bar{q}'(\rho). \quad (36)$$

Notice that $\omega'(q) \geq 0$ when $\bar{q}'(\rho) > 0$ (and $\omega'(q) \leq 0$ if $\bar{q}'(\rho) < 0$), as a result of more (less) high skilled workers being attracted to the financial sector; therefore, the (second order) effect on ω does not change the sign in the case that $\bar{q}'(\rho) > 0$. Can it be that both $\omega'(q) < 0$ and $\bar{q}'(\rho) < 0$? No, since all the terms in the left-hand side of (36) are positive when $\omega'(q) < 0$.

D Numerical solution

We describe first the numerical procedure used to solve Problem (17) for exogenous outside values q and \bar{q} and for exogenous probability of offers ρ . We will then describe how these variables are determined in the steady state equilibrium.

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

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

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

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

$$\beta \left[v(q(1)) + \left(\frac{\mathcal{B}}{1+\lambda} \right) (\mu + \chi + (1-\rho)\gamma(1)) \right] p = -\beta y_\lambda(\lambda) + \frac{\alpha(\gamma + \chi)}{1-\lambda} \quad (37)$$

$$c = \gamma \quad (38)$$

$$g(\lambda, \varepsilon) \psi(q(\varepsilon)) = \mu + \chi + \gamma(\varepsilon) \quad (39)$$

$$v = \beta y(\lambda) - c + \beta \sum_{\varepsilon} g(\lambda, \varepsilon) \Psi(q(\varepsilon)) p(\varepsilon) \quad (40)$$

$$q = \ln(c) + \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left(\mathcal{B} \ln(g(\lambda, \varepsilon)) + q(\varepsilon) \right) p(\varepsilon) \quad (41)$$

$$\begin{aligned} \chi \left\{ \alpha \ln(1-\lambda) + \beta \sum_{\varepsilon} \left[q(\varepsilon) + \mathcal{B} \ln(g(\lambda, \varepsilon)) \right] p(\varepsilon) \right. \\ \left. - \alpha \ln(1-\hat{\lambda}) - \beta \sum_{\varepsilon} \left[(1-\rho)\underline{q} + \rho\bar{q} + \rho\mathcal{B} \ln(g(\hat{\lambda}, \varepsilon)) \right] p(\varepsilon) \right\} = 0 \quad (42) \end{aligned}$$

$$\gamma(\varepsilon) \left[q(\varepsilon) - (1-\rho)\underline{q} - \rho\bar{q} + (1-\rho)\mathcal{B} \ln(g(\lambda, \varepsilon)) \right] = 0 \quad (43)$$

The first three equations are the first order conditions with respect to λ , c , $q(\varepsilon)$, respectively. Equation (40) defines the value for the investor and equation (41) is the promise-keeping constraint. Equations (42) and (43) formalize the Kuhn-Tucker conditions for the incentive-compatibility and enforcement constraints.

Notice that equations (42) and (43) must be satisfied for all values of ε which can take two values. Therefore, we have a system of 9 equations in 9 unknowns: λ , c , v , μ , χ , $q(\varepsilon)$, $\gamma(\varepsilon)$. Once we have solved for the unknowns we can update the functions $\psi(q)$ and $\Psi(q)$ using the solutions for v and μ .

Solving for the steady state equilibrium. The iteration starts by guessing the steady state values of \bar{q} and ρ . Given these two values, we can determine \underline{q} using equation

(26). With these guesses we can solve for the optimal contract as described above. This returns the functions $\mu = \psi(q)$ and $v = \Psi(q)$ in addition to $\lambda = \varphi^\lambda(q)$ and $q(\varepsilon) = \varphi^q(q, \varepsilon)$.

Once we have these functions we determine the new values of \bar{q} and ρ using the free-entry condition (33) and the bargaining condition (34). 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).

E Derivation of inequality index

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

$$\bar{h} = \omega \sum_{j=0}^{\infty} (1 - \omega)^j E_j h_j, \quad (44)$$

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

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

We now turn to the variance which is calculated as

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

This can be rewritten as

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

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

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

To compute the inequality index we simply divide the variance by \bar{h}^2 , where \bar{h} is given by (44). This returns the inequality index (27).

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

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

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

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

Dividing by \bar{h}^2 using the expression for \bar{h} derived in (44), we are able to write the inequality index as

$$\begin{aligned} \text{Inequality index} = & \left(\frac{[1 - (1-\omega) E g(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1-\omega) E g(\hat{\lambda}, \varepsilon)^2]} - \frac{[1 - (1-\omega) E g(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1-\omega) (E g(\hat{\lambda}, \varepsilon))^2]} \right) + \\ & \left(\frac{[1 - (1-\omega) E g(\hat{\lambda}, \varepsilon)]^2}{\omega [1 - (1-\omega) (E g(\hat{\lambda}, \varepsilon))^2]} - 1 \right) \end{aligned} \quad (45)$$

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

F Calibration

We use the 10 moments reported in the bottom section of Table 1 to calibrate 10 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 $\beta = (1-\omega)\hat{\beta} = 0.9375$.
- N is pinned down by the fraction of skilled workers. Since the number of skilled workers is normalized to 1, the fraction of skilled workers is $1/(1+N) = 0.25$.
- ψ is pinned down by the employment share in the financial sector together with the job finding rate in the sector, the probability ρ . Denoting S the number of skilled workers employed in the nonfinancial sector. This can be determined by the employment share in the financial sector, which is equal to $(1-S)/(1+N)$. Next, denote by U the number of skilled workers with managerial ability, currently

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 skilled workers with managerial ability is $U_{t+1} = U_t(1 - \omega)(1 - \rho) + \omega\psi$. After imposing steady state conditions, the two flow equations can be solved for

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

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

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

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

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

- α is pinned down by the time spent innovating. In the model with double-sided limited commitment this maximizes the outside value of the manager and it is determined by the first order condition (16), that is, $\alpha/(1 - \hat{\lambda}) = \rho\beta\mathcal{B}p/(1 + \hat{\lambda})$.
- ν is pinned down by the skill premium in the nonfinancial sector. The earnings of unskilled workers are equal to $w^N = \nu(N/h_0S)^{\nu-1}$ and the earnings of skilled workers in the nonfinancial sector are $w^S h_0 = (1 - \nu)(N/h_0S)^\nu h_0$. Therefore, the skill premium is equal to $w^S h_0/w^N = [(1 - \nu)/\nu]N/S$. Since N has already been calibrated and S has been determined above, this equation determines ν .
- h_0 is pinned down by the share of value added in the financial sector. First, in Section 3 we have derived the average human capital which is equal to

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

The output produced in the financial sector is $(1 - S)\bar{h}(1 - \hat{\lambda}^2)$ and the output produced in the nonfinancial sector is $N^\nu(h_0S)^{1-\nu}$. We can then determine h_0 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.
- 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.
- 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.

- Quadrini, Vincenzo. 2004. "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard", *Journal of Monetary Economics*, 51(4), 713-751.
- Saez, Emmanuel and Thomas Piketty, "income Inequality in the United States - 1913-2002" *Quarterly Journal of Economics*. 118(1), 2003, 1-39.
- Shi, Shouyong. 2006. "Search Theory; Current Perspectives", *Palgrave Dictionary of Economics*.
- Spear, Stephen and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting", *Review of Economic Studies*, 54(4), 599-617.
- Wang, Cheng. 1997. "Incentives, CEO Compensation, and Shareholder Wealth in a Dynamic Agency Model. *Journal of Economic Theory*, 76(1), 72-105.