Risky Investments with Limited Commitment*

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

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

Vincenzo Quadrini  
University of Southern California and CEPR

March 15, 2014

PRELIMINARY AND INCOMPLETE

Abstract

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

*We would like to thank Youssef Cassis, Robert Lucas, Facundo Piguillem, Kim Schoenholtz and other seminar participants at Banca d’Italia, Banque de France, Bank of Turkey, Tel Aviv University, Hebrew University of Jerusalem, New York University, et al., for their comments on previous versions of this work.
1 Introduction

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

The increase in size and importance of the financial sector in the US economy is documented in Phillipon (2008) and Phillipon and Resheff (2009). Figure 1 shows that the GDP share of the financial industry doubled in size between 1970 and 2011. The share of employment has also increased but by less than the contribution to GDP. This is especially noticeable starting in the mid 1980s when the share of employment stopped growing while the share of value added continued to expand. Accordingly, we observe a significant increase in productivity compared to the remaining sectors of the economy. A similar pattern is observed in other countries as shown by Phillipon and Resheff (2013).

The increase in size was associated with a sharp increase in compensation. Clementi and Cooley (2009) show that between 1993 and 2006 the average compensation levels of CEOs in the financial sector increased from parity with other sectors of the economy to nearly double. At the same time compensation of managers became more unequal in the financial sector. Figure 2 shows that the income concentration among managerial occupations in the financial sector (measured by the income share of the top 5%) has

![Size of Finance and Insurance](image-url)

Figure 1: Share of Value Added and Employment

The increase in size was associated with a sharp increase in compensation. Clementi and Cooley (2009) show that between 1993 and 2006 the average compensation levels of CEOs in the financial sector increased from parity with other sectors of the economy to nearly double. At the same time compensation of managers became more unequal in the financial sector. Figure 2 shows that the income concentration among managerial occupations in the financial sector (measured by the income share of the top 5%) has
increased significantly compared to the rest of the economy.

![Income Share of Top 5%](image)

**Figure 2: Share of the top 5%**

Although productivity in the financial sector has increased compared to the other sectors of the economy, the valuation of financial companies has not increased as much as nonfinancial companies. Figure 3 plots the average ratio of market to book value of equity for publicly listed financial and nonfinancial firms. Starting in the early 1980’s, the market valuation of financial firms displays a flat trend while the valuation of nonfinancial firms has continued to grow. This may be a reflection of compensation practices in a sector where managers were able to retain a larger share of the surplus.

The changes described above took place during a period in which the organizational structure of the financial sector was also changing, with traditional partnerships replaced by public companies. Until 1970 the New York Stock Exchange prohibited member firms from being public companies. When the organizational restriction on financial companies was relaxed, there was a movement to go public and partnerships began to disappear. Merrill Lynch went public in 1971, followed by Bear Stearns in 1985, Morgan Stanley in 1985, Lehman Brothers in 1994 and Goldman Sachs in 1999. Other venerable investment banks were taken public and either absorbed by commercial banks or converted to bank holding companies. Today, there are very few partnerships remaining and they are small. The same evolution occurred in Britain where the closed ownership Merchant Banks virtually disappeared.

The transition from partnerships to public companies had two important implications. The first implication was to increase competition for managers in the financial sector. The second was to alter the structure of contractual arrangements between investors and managers in ways that weakened commitments.¹

¹The transition from partnership to public corporations also had some implications for the liability
Why did the transformation from partnerships to corporations increased competition for managerial talent? As financial firms became public, they had greater access to capital (through the sales of shares) which facilitated their growth. But capital is only one of the production factors. Human capital is also important. Therefore, as more financial capital was coming in, more managerial capital was needed and this increased the competition (demand) for managers, in particular, and workers in general.\(^2\)

Why did the transformation from partnerships to corporations weaken the commitment of investors and managers? Many argued that a partnership was a preferred form of organization for investment firms because managers and investors were the same people and it was the partners own assets that were at risk. Public companies, on the other hand, are organizations with significant separation between ownership (shareholders) and investment control (managers), and it is well understood that they are characterized by significant agency problems.\(^3\) In a world where contracts are fully enforceable, agency problems are solved with the optimal design of contracts. In reality, however, enforce-

\(^2\)Roy Smith, a former partner at Goldman Sachs described the evolution of the relationship between compensation and firm structure as follows: “In time there was an erosion of the simple principles of the partnership days. Compensation for top managers followed the trend into excess set by other public companies. Competition for talent made recruitment and retention more difficult and thus tilted negotiating power further in favor of stars. You had to pay everyone well because you never knew what next year would bring, and because there was always someone trying to poach your best trained people, whom you didn’t want to lose even if they were not superstars. Consequently, bonuses in general became more automatic and less tied to superior performance. Compensation became the industry’s largest expense, accounting for about 50% of net revenues”, Wall Street Journal February 7, 2009.

\(^3\)This is largely consistent with the literature on incomplete contracts. According to Grossman and Hart (1986) and Hart and Moore (1990), more efficient organizational forms are those where the agents who control the investment surplus own a larger share of the assets.
ment is limited and the transformation of partnerships to public companies has further reduced enforcement. On the one hand, shareholders could replace managers and renege future promises made to them. On the other, managers have some discretion in the operational decisions of the firm and could leave. The organizational change also increased the mobility of managers. In a partnership, the ownership shares were relatively illiquid so it was difficult for partners to liquidate their ownership positions and move to other firms. Also important was the process of becoming a partner. In the typical firm, new professionals are hired as associates and, after a trial period, they are either chosen to be partners or released. In this environment separation is viewed as a signal of inferior performance, thus affecting the external option of a financial professional. Becoming a partner, on the other hand, represents a firm commitment to continued employment on the part of the other partners.

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

We address this question by developing a model in which investors compete for and hire managers to run investment projects, with each investor-manager pair representing a financial firm. Two features of the model are especially important. The first feature is that production depends on the human capital of the manager which can be enhanced, within the firm, with costly investment. Human capital accumulation can be understood as acquiring new skills by engaging in risky financial innovations (e.g. implementing new financial instruments which may or may not have positive returns). The second important feature is that human capital can be transferred outside the firm by managers. This generates a conflict of interest between investors and managers: while the interest of investors is for the value of human capital inside the firm, managers also care about the outside value. As a result, the investment desired by investors may be different than the investment desired by managers. Then, if investors cannot control the firm policies either directly or indirectly through a credible compensation scheme, managers may deviate from the optimal policies.

We first characterize the optimal contract with one-sided limited commitment. In this environment only the investor commits to the contract. We interpret this case as capturing the economic environment that prevailed in the period preceding the change in organizational structure (from partnerships to public companies). Although in this period there was not a clear separation between ownership and management, still, partners

---

4The New York Stock Exchange regulatory change mentioned above has been an important factor allowing financial corporations to become public companies. However, this does not tell us why they have chosen to do so. In several cases firms were simply acquired by public companies but in others it was an important strategic decision. Charles Ellis (2008) in his history of Goldman Sachs—the last major firm to go public—suggests that the major motive for financial partnerships to become public was to increase capital for their proprietary trading operations through an IPO. The goal of this paper is not to understand why financial companies have chosen to become public. Rather, we want to understand the consequences of having a financial industry changing from a partnership type of organization to public companies.
could quit the partnership, which motivates our choice of one-sided limited commitment to characterize the contractual relationships in the earlier period. After studying the environment with one-sided limited commitment, we analyze the optimal contract with double-sided limited commitment. In this environment, contracts are not fully enforceable for both managers and investors. In particular, investors can renege on promises made in the past and could replace the managers. We interpret this case as representative of the most recent period characterized by a clearer separation between ownership and management: When investors (shareholders) are different from managers, their commitment becomes an issue.

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

To make the outside value of managers endogenous and to study the implications for the whole economy, we embed the micro structure in a general equilibrium model with two sectors, where the financial sector produces intermediate services for the nonfinancial sector. In the general model we formalize the increased competition for managers by lowering the cost to create jobs in the financial sector while the weakened commitment is captured by the shift to a regime with double-sided limited commitment. We then show that these structural changes can generate (i) greater risk-taking in both sectors; (ii) larger share and higher relative productivity of the financial sector; (iii) greater income inequality within and between sectors; and (iv) lower valuation of financial companies. An important feature of the general model is that, even if the structural changes take place only in the financial sector, general equilibrium effects induce higher risk-taking and higher income concentration also in the nonfinancial sector.

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

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

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 market conditions for managers, it becomes important to derive these conditions endogenously. A second assumption typically made in principal-agent models is that investors fully commit to the contract. However, the clearer separation between investors and managers that followed the transformation of financial partnerships to public companies and the associated competition for managerial talent, could have also reduced the commitment of investors.

In this paper we relax both assumptions: we make the outside option of managers endogenous and we allow for the limited commitment of investors. As we will see, the relaxation of both assumptions are key for the central results of the paper.\footnote{Cooley, Marimon and Quadrini (2004) also endogenized the outside value of entrepreneurs but kept the assumption that investors commit to the long-term contract. Marimon and Quadrini (2011) relaxed both assumptions and, using a model without uncertainty, showed that differences in “barriers to competition”, can result in income differences across countries. In these two papers, however, uncertainty does not play a significant role while it is central to the analysis of the current paper. Furthermore, the current model features two sectors (financial and nonfinancial) and shows that changes in one sector can have important effects, through the general equilibrium, on the other sector.}

The empirical facts described in the introduction have also motivated other studies. The models used in these studies can capture some of the empirical facts but we are not aware of models that can capture all of them simultaneously. We are also unaware of any study that connects changes in the organizational structure of the financial sector with the increased competition for managerial talent. Cheng, Hong and Scheinkman (2012) and, in a general equilibrium Edmans and Gabaix (2011), explain how in a Principal-Agent relationship with a fixed sharing rule, an \textit{exogenous} increase in risk can result in higher compensation for risk-averse financial managers in order to satisfy their participation and incentive constraints. Bolton, Santos and Scheinkman (2012) argue that it is “cream skimming” in the more opaque financial transactions—those taking place in over-the-counter or bespoke markets—that have encouraged excessive compensation of financial managers and the excessive large share of GDP of the financial services industry. In our paper, instead, we propose a model that could generate the empirical facts as a consequence of the organizational change that has taken place in the financial sector during the last three decades. In our model, instead, the increase in risk is generated endogenously as a consequence of greater \textit{competition} and weaker \textit{commitment}. Through
general equilibrium effects, this can generate higher risk also in the nonfinancial sector even even there are no structural changes in this sector.

2 The model

We start with the description of the financial sector and the contracting relationships that are at the core of the model. After the characterization of the financial sector, we will embed it in a general equilibrium in Section 5.

Managers are skilled workers employed by a firm who have the ability to produce and develop innovative projects. But managers could be mobile and when they choose to leave the firm, at least part of the know-how created with innovative projects can be transferred by them to other firms. The employment relationship is regulated by a contract between an investor—the owner of the firm—and a manager.

In the case of a partnership we should think of the investor as the representative of all partners, who are also the managers of the firm. Effectively, each individual partner enters into a contractual relationship with all other partners and they are represented by a fictitious ‘investor’. To simplify the analysis we assume that a partnership is composed of a large number of partners so that the risk induced by the action of an individual partner is negligible for the whole partnership. By further assuming that the income of the firm is linear in the number and skills of partners, we can characterize the optimal contract by focusing on the relationship between a single risk-neutral investor and a single risk-averse manager. Although this assumption may appear a major oversimplification, it allows us to capture some of the key differences between partnerships and public company without losing tractability.

Because in a partnership the investor is the representative of the partners, it is unlikely that the investor reneges on promises made to the partners he represents. Therefore, to characterize the contractual relationship in a partnership, we assume one-sided limited commitment where the investor commits to the contract but the manager does not commit. In the case of public companies, instead, investors are the shareholders of the company, distinct from managers. The separation between ownership and management makes the possibility of reneging on previous promises to managers highly relevant. Of course, if it were possible to write formal contracts in which future promises are legally binding, shareholders would not be able to renege on these promises. However, making all promises legally binding may not be feasible or desirable. Based on this premise, we characterize the contractual relationship in a public company with double-sided limited commitment where there is limited commitment also for investors.

To summarize, what distinguishes a partnership from a public company in our model is the degree to which contracts are enforceable: one-side limited commitment for partnerships and double-sided limited commitment for public companies.

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

Investors (representative of partners or shareholders) are risk-neutral agents who are the residual claimants to the income of the firm. The expected lifetime utility is

\[ V_0 = E_0 \sum_{t=0}^{\infty} \beta^t (\Pi_t - C_t), \]

where \( \Pi_t \) denotes the cash flows generated by the firm and \( C_t \) is the compensation of managers. Thus, \( \Pi_t - C_t \) represents the payment received by investors.

Managers are risk averse with expected lifetime utility,

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

where \( C_t \) is the manager’s compensation (consumption) and \( \lambda_t \) is the effort needed to innovate as described below. The period utility satisfies \( u'(>0), u''(<0) \) and \( e(0) = 0, e'(>0), e''(>0). \)

Managers are characterized by human capital \( h_t \). The income generated by the firm is equal to \( Ph_t \). At this stage the variable \( P \) is simply a constant. When we extend the model to a general equilibrium this variable is endogenously determined as the price of financial services in terms of final goods.

Human capital can be increased by investing \( I_t = \kappa(\lambda_t)h_t \). The investment generates a new implementable project \( i_{t+1} \) according to the technology,

\[ i_{t+1} = \lambda_t h_t \varepsilon_{t+1}. \]

The variable \( \lambda_t \) determines the scale of the investment. Larger scales are associated with higher cost for the firm and require more effort from the manager. The variable \( \varepsilon_{t+1} \in \{\bar{\varepsilon}, \bar{\varepsilon}\} \) is an i.i.d. stochastic variable, capturing the innovation risk. The function \( \kappa(.) \) is strictly increasing, strictly convex and satisfies \( \kappa(0) = 0 \).

We think of \( i_{t+1} \) as a new project that enhances the human capital of the manager only if the project is implemented in a firm—either the current or new firm. With its actual implementation, the human capital of the manager becomes \( h_{t+1} = h_t + i_{t+1} \). If the new project \( i_{t+1} \) is not implemented in a firm—for instance, if the manager leaves the financial sector and finds occupation outside the financial industry—her human capital remains \( h_t \). Therefore, if a new project is implemented after the development stage, it becomes embedded human capital. Otherwise it fully depreciates. The importance of this assumption will become clear later.\(^6\)

\(^6\)The assumption that the pre-existing human capital does not depreciate when the manager leaves the financial industry is not essential for the qualitative properties of the model. It is only made to maintain the linear homogeneity in \( h_t \). The alternative assumption that the whole human capital depreciates when the manager leaves the financial sector would lead to similar qualitative properties. However, we would lose the linear homogeneity property in \( h_t \). As we will see, this property allows us to work with a representative firm even if firms employ managers with different \( h_t \).
To use a compact notation, we denote by $\pi(\lambda_t) = P - \kappa(\lambda_t)$ the income, net of the investment cost, generated by the firm per each unit of managerial human capital. Furthermore, we denote by $g(\lambda_t, \varepsilon_{t+1}) = 1 + \lambda_t \varepsilon_{t+1}$ the gross growth rate of human capital, provided the manager remains employed in a financial firm. Then, the firm’s cash flow and the evolution of human capital can be written as

$$\Pi_t = \pi(\lambda_t) h_t, \quad (1)$$

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

**Information structure and timing.** All variables are public information with the exception of $\lambda_t$, the innovation scale chosen by the manager. Although this variable is not observable, the investor can infer the value of $\lambda_t$ by observing the investment $I_t = \kappa(\lambda_t) h_t$ since $h_t$ is public information. However, we assume that the investor observes $I_t$ (and indirectly $\lambda_t$) at the end of the period, after the compensation of the manager in the current period. This assumption implies that current compensation $C_t$ cannot be conditional on the current choice of $\lambda_t$. It can only be contingent on past innovations. An alternative assumption is that, given the choice of $\lambda_t$, the investment cost is incurred at $t + 1$. Therefore, the firm has to wait next period before being able to infer $\lambda_t$. This alternative assumption would not change the properties of the model but would require a more complex notation.

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

Denote by $Q_{t+1}(h_t)$ the manager’s outside value at the beginning of period $t + 1$ without an external offer and by $\overline{Q}_{t+1}(h_{t+1})$ the outside value with an offer. The expected outside value at $t + 1$ of a manager with previous human capital $h_t$ is equal to

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

For the moment we take the probability $\rho$ and the outside value functions $Q_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ as exogenous. At this stage we only assume that $Q_{t+1}(h_t)$ and $\overline{Q}_{t+1}(h_{t+1})$ are strictly increasing and differentiable. However, when we extend the model to a general equilibrium in Section 5, these terms will be derived endogenously. This is an important innovation that will be central for the derivation of some of the key results of this paper.

In addition to having the ability to quit, the manager has full control over $\lambda_t$. Full control is allowed by the assumption that this variable is not directly observable by investors. The investor can only infer the value of $\lambda_t$ at the end of the period, after the payment of the manager compensation in the current period. This implies that, in
absence of proper incentives, the value $\lambda_t$ chosen by the manager may not be efficient. In particular, the manager may be tempted to increase $\lambda_t$ in order to raise the outside value because it does not take into account in investment cost $I_t$ incurred by the firm. The manager only internalizes the dis-utility cost $e(\lambda_t)$. Therefore, there are two sources of frictions in the decision problem of the manager: the ability to quit and discretion in the choice of $\lambda_t$.

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

In a public company, instead, investors are the shareholders of firms and they are distinct from managers. Because of this separation, the possibility of reneging on previous promises could become central to the contractual relationship between shareholders and managers of a public company. To capture this possibility, we assume that in a public company there is double-sided limited commitment: managers could quit and investors could renege on future promises.

**Definition of a contract.** We conclude the description of the model with a formal definition of a contract.

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

Notice that the payment made to the manager in period $t$ is not conditional on $\lambda_t$ but only on past values. This is because $\lambda_t$ become public information only after the payment of $C_t$.

**3 Optimal contract**

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

3.1 One-sided limited commitment: The case of partnerships

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

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

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

which must be satisfied for all \( t \geq 1 \). Notice that the contract starts at time zero but the constraint must be satisfied starting at \( t \geq 1 \).

A second constraint takes into account that the manager has full control in the choice of effort and could deviate from the \( \lambda_t \) recommended by the planner (incentive-compatibility). Denote by \( \hat{\lambda}_t \) the innovation effort chosen by the manager when she deviates from the recommended innovation. By deviating, the manager anticipates leaving the firm at the beginning of the next period. Therefore, \( \hat{\lambda}_t \) maximizes the anticipated value of quitting, that is,

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

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

The optimal deviation \( \hat{\lambda}_t \) can be characterized with the first order condition,

\[ e_1(\hat{\lambda}_t) = \beta E_t D_2 \left(h_t, g(\hat{\lambda}_t, \varepsilon_{t+1}), h_t, \rho \right) g_1(\hat{\lambda}_t, \varepsilon_{t+1}) h_t. \]  

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

An important feature of the optimal deviations $\hat{\ell}_t$ and $\hat{\lambda}_t$ is that they are not affected by current compensation $C_t$ because of the assumption that $\lambda_t$ becomes public information only after paying the manager’s compensation $C_t$. The manager can still be punished at $t+1$ by cutting $C_{t+1}$. However, at that point, the ability to quit sets a lower bound to the feasible punishment. If $C_t$ could be conditioned on $\hat{\lambda}_t$, investors could punish managers’ deviation by reducing $C_t$. By further assuming that the utility function satisfies $u(0) = -\infty$, the planner would have unlimited power to punish managers and, de-facto, they would not have discretion in the choice of $\lambda_t$.

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

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

(7)

The left-hand-side is the value that the manager receives if she chooses the innovation recommended by the planner, $\lambda_t$. The right-hand-side is the value achieved by deviating from the recommended policy, that is, when the manager chooses $\hat{\lambda}_t$ as determined in (5).

As observed above, current compensation $C_t$ cannot be contingent on the actual choice of $\lambda_t$ since this variable becomes public information after the payment of $C_t$. Therefore, the current utility from consumption is the same with or without deviation.

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

$$
\max_{\{C_t, \lambda_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left( \pi(\lambda_t) h_t - C_t \right) + \hat{\mu}_0 \sum_{t=0}^{\infty} \beta^t \left( u(C_t) - e(\lambda_t) \right) \right\} \tag{8}
$$

s.t. (2), (4), (7).

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

12
Following Marcet and Marimon (2011), the problem can be written recursively as

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

(9)

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

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

The variable \( \tilde{\mu}_t \) captures the value of the contract for the manager. It evolves over time according to \( \tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1}) \). Therefore, the value for the manager increases any time the incentive-compatibility constraint or the enforcement constraint are binding. Higher values imply higher compensation (current and future) that is necessary to prevent the manager from deviating and quitting.

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

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

(10)

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

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

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

$$
\left( \frac{\tilde{\mu}_t + \tilde{\chi}_t}{h_t} \right) e_1(\ell_t + \lambda_t) - \pi_1(\lambda_t) = \beta E_t \left[ W_1(h_{t+1}, \tilde{\mu}_{t+1}) - \left( \tilde{\chi}_t + \tilde{\gamma}_t(\varepsilon_{t+1}) \right) D_2(h_t, h_{t+1}, \rho) \right] g_1(h_t, \varepsilon_{t+1}).
$$

(11)

The left-hand side is the marginal cost of innovation per unit of human capital. This term is increasing in \( \lambda_t, \tilde{\mu}_t \) and \( \tilde{\chi}_t \). The right-hand-side is the expected marginal benefit from investing, net of participation costs. Binding incentive-compatibility and enforcement constraints imply positive values of \( \tilde{\chi}_t \) and \( \tilde{\gamma}_t(\varepsilon_{t+1}) \) which tend to reduce the right-hand-side of equation (11). Therefore, to the extent that the right-hand-side
term declines with $\tilde{\chi}_t$ and $\tilde{\gamma}_t(\varepsilon_{t+1})$, we have that binding constraints reduce innovation $\lambda_t$.

Intuitively, to retain the manager, the value of staying must increase or the value of quitting must decline. The value of staying can be increased by promising higher compensation and by requiring lower effort (that is, by changing $\lambda_t$). The value of quitting can be reduced by choosing a lower $\lambda_t$. Both channels lead to a decrease in $\lambda_t$ when the constraints become binding.

We are particularly interested in understanding how higher competition (captured by a higher value of $\rho$) affects the optimal investment policy. In an economy with higher $\rho$ managers have better outside opportunities, implying that the initial $\mu_0$ is higher (given $h_0$). We obtain the following result which we prove formally in Appendix A).

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

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

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

In the environment with double-sided limited commitment, which we think represents the contractual environment in public companies, managers are free to leave the firm and investors can renegade promises made to managers. This implies that the contract is renegotiated whenever the value for the manager exceeds her outside value. For simplicity, we assume that in case of renegotiation the investor has the whole bargaining power. Therefore, the value of the manager is renegotiated down to her outside value. Based on this, the planner also faces the constraint that the value of the contract for the manager cannot exceed the outside value.

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

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

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

(12)

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

The contract simply prescribes a consumption plan which is determined by (10) with $\tilde{\mu}' = \tilde{\gamma}(\varepsilon')$ and the innovation $\hat{\lambda}$ solves the first order condition (6). Since $D_{2,3} > 0$, an increase in competition captured by the parameter $\rho$ increases the right-hand-side of (6), that is, it increases the marginal benefit of innovation for the manager. This is stated formally in the next proposition.

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

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

## 4 Normalization with log-utility

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

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

$$u(C) - e(\lambda) = \ln(C_t) - e(\lambda_t),$$

$$Q_{t+1}(h_t) = q + B \ln(h_t),$$

$$\overline{Q}_{t+1}(h_{t+1}) = \bar{q} + B \ln(h_{t+1}),$$

where $q$, $\bar{q}$ and $B \equiv \frac{1}{1-\beta}$ are constant.
Although the functional forms for the outside values may seem arbitrary, we will see that in the general equilibrium they take exactly these forms.

We start by normalizing the value of the contract for the investor which can be expressed recursively as
\[ V_t = \pi(\lambda_t)h_t - C_t + \beta E_t V_{t+1}. \]
This can be rewritten as,
\[ v_t = \pi(\lambda_t) - c_t + \beta E_t g(\lambda_t, \varepsilon_{t+1})v_{t+1}, \quad (13) \]
where \( v_t = V_t/h_t \) and \( c_t = C_t/h_t \).

The value of the contract for a manager can be expressed recursively as
\[ Q_t = \ln(C_t) - e(\lambda_t) + \beta E_t Q_{t+1}. \]
Defining \( q_t = Q_t - B \ln(h_t) \), we can rewrite it in normalized form as,
\[ q_t = \ln(c_t) - e(\lambda_t) + \beta E_t \left[ B \ln g(\lambda_t, \varepsilon_{t+1}) + q_{t+1} \right]. \quad (14) \]

Next we consider the enforcement constraint after the realization of \( \varepsilon_{t+1} \),
\[ Q_{t+1}(h_{t+1}) \geq (1 - \rho) \cdot Q_{t+1}(h_t) + \rho \cdot Q_{t+1}(h_{t+1}). \]

Using \( q_{t+1} = Q_{t+1}(h_{t+1}) - B \ln(h_{t+1}) \) and the functional forms specified in Assumption 1, the enforcement constraint (7) can be rewritten as
\[ q_{t+1} \geq (1 - \rho)q + \rho \bar{q} - (1 - \rho)B \ln g(\lambda_t, \varepsilon_{t+1}). \quad (15) \]

The right-hand-side depends on \( \lambda_t \) (provided that \( \rho < 1 \)). Thus, investment affects the outside value of the manager and, when the enforcement constraint is binding, it affects compensation. This property is a direct consequence of the assumption that the outside value of the manager without an external offer depends on \( h_t \), while the outside value with an external offer depends on \( h_{t+1} \). If both values were dependent on the embedded human capital \( h_{t+1} \), the last term in (15) would disappear. The value of quitting would still depend on \( \rho \) but it would not affect the choice of \( \lambda_t \).

The constraint that insures that the manager chooses the optimal allocation of effort, incentive-compatibility, is
\[ -e(\lambda_t) + \beta E_t Q_{t+1}(g(\lambda_t, \varepsilon_{t+1})h_t) \geq -e(\hat{\lambda}_t) + \beta E_t \left[ (1 - \rho) \cdot Q_{t+1}(h_t) + \rho \cdot Q_{t+1}(g(\hat{\lambda}_t, \varepsilon_{t+1})h_t) \right], \]
where \( \lambda_t \) is the investment recommended by the contract and \( \hat{\lambda}_t \) is the investment chosen by the manager under the assumption that she will quit at the beginning of next period. After normalizing, the incentive-compatibility constraint becomes,
\[ -e(\lambda_t) + \beta E_t \left[ q_{t+1} + B \ln g(\lambda_t, \varepsilon_{t+1}) \right] \geq -e(\hat{\lambda}_t) + \beta E_t \left[ (1 - \rho)q + \rho \bar{q} + \rho B \ln g(\hat{\lambda}_t, \varepsilon_{t+1}) \right]. \quad (16) \]
We can now provide a more explicit characterization of the manager’s optimal deviation $\hat{\lambda}_t$. Using $g(\lambda, \varepsilon) = 1 + \lambda \varepsilon$, condition (6) can be written as

$$e_1(\hat{\lambda}_t) = \rho \beta BE_t \left( \frac{\varepsilon_{t+1}}{1 + \hat{\lambda}_t \varepsilon_{t+1}} \right).$$  \hspace{1cm} (17)

We can now see more explicitly that $\hat{\lambda}$ increases with $\rho$, as stated more generally in Proposition 2. Therefore, when the manager faces better outside options, the strategic incentive to innovate increases.

**One-sided limited commitment: The case of partnerships.** The original contractual problem (8) with one-sided limited commitment can be reformulated in normalized form using the ‘promised utility’ approach: This maximizes the normalized investor’s value subject to the normalized promise-keeping, limited enforcement and incentive-compatibility constraints, that is,

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

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

The solution provides the innovation 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 the innovation policy $\lambda$ and the initial human capital $h$, we can reconstruct the whole sequence of human capital through the law of motion $h' = g(\lambda, \varepsilon')$. From this we are able to reconstruct the original, non-normalized variables $C = ch$, $Q = q + B \ln(h)$ and $V = vh$.

Since there is a one-to-one mapping from the normalized policies to the original (non-normalized) variables, we can characterize the optimal contract by focusing on the normalized policies, which satisfy the first order conditions

$$c = \mu,$$

$$\mu(\varepsilon') = \frac{\mu + \chi + \gamma(\varepsilon')}{g(\lambda, \varepsilon')}.$$  \hspace{1cm} (21)

The variables $\mu$, $\gamma(\varepsilon')$ and $\chi$ are the Lagrange multipliers for constraints (14)-(16). These multipliers are related to the multipliers used in Section 3 as $\mu = \bar{\mu}/h$, $\gamma(\varepsilon')/h$ and $\bar{\chi}/h$. The detailed derivation of the first order conditions is provided in Appendix B.
Double-sided limited commitment: The case of public companies. With double-sided limited commitment, investors renegotiate promises that exceed the outside value of managers. Thus, the value of the contract for the manager is always equal to the outside value, that is, the enforcement constraint is always binding. Anticipating renegotiation, the best strategy for the manager is to choose the innovation that maximizes the outside value \( \hat{\lambda} \). This is determined by condition (17). Problem (12) can then be reformulated in normalized form as,

\[
v(q) = \max_{c,q(\varepsilon)} \left\{ \pi(\hat{\lambda}) - c + \beta Eg(\hat{\lambda}, \varepsilon)v(q(\varepsilon)) \right\}
\]

subject to

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

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

Problem (22) is a special case of problem (18) where we have replaced the incentive-compatibility constraint (16) with \( \lambda = \hat{\lambda} \), and the enforcement constraint (15) is always binding. Notice that the decision variables \( c \) and \( q(\varepsilon) \) are fully determined by the promise-keeping and enforcement constraints. Therefore, the problem can be solved without performing any optimization, besides solving for \( \hat{\lambda} \).

4.1 Contract properties

In this subsection we illustrate the properties of the optimal contract numerically. The specific parameter values will be described in Section 6.1 where we conduct a quantitative analysis with the general model. The computational procedure used to solve the optimal contract is described in Appendix D\(^7\).

The solution to the contractual problems (18) and (22) provide the optimal policies for investment, \( \lambda = \varphi^\lambda(q) \), manager’s consumption, \( c = \varphi^c(q) \), and continuation utilities, \( q(\varepsilon) = \varphi^q(q, \varepsilon) \). Because of the normalization, these policies are independent of \( h \).

However, once we know the normalized policies and the initial human capital \( h_0 \), we can construct the whole sequences of human capital \( h \), non-normalized consumption, \( C = ch \), and non-normalized lifetime utility, \( Q = q + B \ln(h) \). Therefore, to characterize the optimal contract we can focus on the normalized policies as characterized by the first order conditions (19)-(21), in the case of one-sided limited commitment, and by the constraints of problem (22) in the case of double-sided limited commitment.

The dynamics of promised utilities. The top panels of Figure 4 plot the values of next period normalized continuation utilities, \( q(\varepsilon) = \varphi^q(q, \varepsilon) \), as functions of current

\(^7\)Without loss of generality, we assume for the rest of the paper that \( \bar{\varepsilon} = 1 \).
normalized utility, $q$, for the environments with one-sided and double-sided limited commitment. We have also plotted the 45 degree line which allows us to see more clearly the dynamics of the contract in response to the shock. If the continuation utility is below (above) the 45 degree line, the next period $q$ is smaller (bigger) than the current $q$. The vertical lines indicate the initial normalized values of the contract for the manager, $\bar{q}$. We have not specified how to determine the initial value of $q$. This will be described when we embed the model in a general equilibrium.

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

We discuss first the case with one-sided limited commitment. The contract starts with an initial $\bar{q}$ indicated by the vertical line. Then, if the investment does not succeed ($\varepsilon = 0$), the next period value of $q$ remains the same. If the investment succeeds ($\varepsilon = 1$), the next period $q$ declines until it reaches a lower bound. At that point the value of $q$ remains constant. It is important to remember, however, that these are normalized utilities. Therefore, the fact that $q$ declines does not necessarily mean that the actual (non-normalized) utility $Q = q + B\ln(h)$ declines. In the event of a positive realization of $\varepsilon$, the non-normalized lifetime utility $Q$ increases but less than $B\ln(h)$. In the case of a negative shock, instead, $Q$ (and $q$) remains constant.

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

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

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

**Investment.** The bottom panels of Figure 4 plot the investment policy $\lambda$. In the environment with one-sided limited commitment, the enforcement constraint is not binding for high values of $q$. As a result, $\lambda$ 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 $\lambda$ increases the outside value for the manager and must be associated with a higher promised utility. Since this is costly for the investor, the optimal $\lambda$ is lower for low values of $q$ (although quantitatively the dependence is very small).

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

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

5 General model

We now embed the financial sector in a general equilibrium, we assume that there are two sectors: financial and nonfinancial. Given the service nature of the financial sector, we assume that its output is used as an intermediate input in the production of final goods in the nonfinancial sector.

There are two types of agents—a unit mass of investors and a unit mass of risk-averse workers, with the same preferences as described earlier. In particular we focus on the case of log-utility for workers as specified in Section 4. However, we now assume that workers die with probability $\omega$. The discount factor used by worker should then be considered the product of the actual intertemporal discount factor $\hat{\beta}$, and the survival probability $1 - \omega$, that is, $\beta = \hat{\beta}(1 - \omega)$. In every period a mass $\omega$ of newborn workers with initial human capital $h_0$ enter the economy. In this way the population size remains constant over time. The motivation for adding this particular demographic structure is to keep the distribution of $h_t$ among living workers stationary. The finite lives of workers together with constant initial human capital $h_0$ for newborn workers guarantee that the distribution of $h_t$ converges to an invariant distribution and aggregate variables are stationary in level.

A fraction $\psi$ of newborn workers have the skills to be employed in the financial sector and can find occupation in that sector with some probability. To simplify the analysis we assume that a worker retains the ability to work in the financial sector only if she continues to be employed in the financial sector. Therefore, when a worker exits the financial sector or a newborn worker with financial skills does not find immediate occupation in the financial sector, she can subsequently find occupation only in the nonfinancial sector. Therefore, new workers that enter the financial sector only come from the pool of newborn workers. These assumptions are not important for the qualitative properties of the model but simplifies the analysis.

5.1 Financial sector

The technology and contractual structure of the financial sector is as described in the previous sections. The only difference is that the output produced by the financial
sector is in the form of intermediate services purchased by nonfinancial firms to produce final goods. The purchasing price, denoted by $P_t$, will be determined in the general equilibrium. The cash flows generated by a financial firm is as specified before, that is, 

$$
\Pi_t = \left( P_t - \kappa(\lambda_t) \right) h_t,
$$

where now $P_t$ is the equilibrium price for financial services. For an individual firm, however, $P_t$ is fixed and the characterization of the optimal contract is analogous to the previous section.

Compared to the basic model, we now place some structure on the hiring process. We assume that the labor market is characterized by matching frictions. Workers in search of an occupation in the financial sector find employment if they are matched with vacancies funded by investors. Consistent with the previous interpretation of a financial firm, in a partnership investors are the representatives of partners while in a public company they are shareholders, separate from managers.

Since workers are heterogeneous in human capital, we assume directed search. Vacancies specify the level of human capital $h$ and the initial value of the contract for the worker $Q_t(h)$. The cost of posting a vacancy is $\tau h$.

Denote by $N_t(h, \overline{Q}_t)$ the number of vacancies posted for workers with human capital $h$ offering lifetime utility $\overline{Q}_t(h)$. Furthermore, denote by $S_t(h, \overline{Q}_t)$ the number of workers with human capital $h$ in search of an occupation in the financial sector with posted value $\overline{Q}_t$. The number of matches is determined by the function $m_t(h, \overline{Q}_t) = AN_t(h, \overline{Q}_t)^\eta S_t(h, \overline{Q}_t)^{1-\eta}$. From the matching function we derive the probabilities that a vacancy is filled, $\phi_t(h, \overline{Q}_t) = m_t(h, \overline{Q}_t)/N_t(h, \overline{Q}_t)$, and the probability that a searching worker finds a financial occupation, $\rho_t(h, \overline{Q}_t) = m_t(h, \overline{Q}_t)/S_t(h, \overline{Q}_t)$. Free entry in the financial sector implies that, for any level of human capital $h$, $\phi_t(h, \overline{Q}_t) V_t(h, \overline{Q}_t) = \tau h$.

We now take advantage of the properties derived earlier in the case of log-utility for workers. We have seen in Section 4 that the value of the contract for the investor is linear in $h$, that is, $V_t(h, \overline{Q}_t) = v_t(\overline{Q}_t) h$, and the normalized value for a newly hired worker can be written as $\overline{q}_t = \overline{Q}_t - B \ln(h_t)$. To determine $\overline{q}_t$, we only need to define a menu of posted contracts for all possible levels of human capital $h$. Focusing on a symmetric equilibrium in which the probability of filling a vacancy is independent of $h$, the free-entry condition in the financial sector can be rewritten in normalized form as

$$
\phi_t(\overline{q}_t) v_t(\overline{q}_t) = \tau. \tag{23}
$$

Appendix C discusses the equilibrium conditions in the labor market in more detail and shows that, when matched, workers receive a fraction $1 - \eta$ of the surplus. This is a standard property in models with directed search.

The next step is to characterize the outside value of workers, which is defined as

$$
D(h_{t-1}, \lambda_{t-1}, \varepsilon_t) = (1 - \rho_t)Q_t(h_{t-1}) + \rho_t Q_t(h_t). \tag{24}
$$

We have already defined the matching probability $\rho_t$ and the value of the match for the worker $Q_t(h_t) = \overline{q}_t - B \ln(h_t)$. What is left to define is the value of not being
matched, \( Q_t(h_{t-1}) \). In this case the worker becomes unemployed and will search for a job in the nonfinancial sector (remember that the worker can no longer be employed in the financial sector once she exits the sector). Therefore, denoting by \( U_t(h_{t-1}) \) the value of being unemployed, we have that \( Q_t(h_{t-1}) = U_t(h_{t-1}) \). In order to define the unemployment value, however, we need to specify the nonfinancial sector.

### 5.2 Nonfinancial sector

Workers in the nonfinancial sectors can also innovate using the same technology as in the financial sector, that is, by investing \( I_t = \kappa(\lambda_t) h_t \), human capital evolves according to

\[ h_{t+1} = (1 + \lambda_t \varepsilon_{t+1}) h_t, \]

The production of final goods depends on the human capital of the worker—as in the financial sector—but it also uses intermediate services produced by the financial sector. Specifically, a worker with human capital \( h_t \) employed in the nonfinancial sector produces final goods \( F(x_t)h_t \), where \( x_t \) denotes the input of intermediate services per unit of human capital. The function \( F(\cdot) \) is strictly increasing and concave. The cash flows generated by a nonfinancial firm is then

\[ \Pi_t = \left( F(x_t) - x_t P_t - \kappa(\lambda) \right) h_t, \]

where \( P_t \) is the price of financial services produced in the financial sector.

Since there are no frictions in the choice of financial services, \( x_t \) is determined by the first order condition \( F_1(x_t) = P_t \). Therefore, the financial input is only a function of the price \( P_t \), which allows us to rewrite the cash flows more compactly as \( \Pi_t = \pi(P_t, \lambda_t) h_t \).

The price \( P_t \) is determined in equilibrium and it is taken as given by an individual firm. The process of job creation, also based on directed search, and the determination of the initial values of the contract are the same as in the financial sector. The outside value for workers employed in the nonfinancial sector is given by

\[ D(h_{t-1}, \lambda_{t-1}, \varepsilon_t) = (1 - \rho_t) U_t(h_{t-1}) + \rho_t Q_t(h_t), \quad (25) \]

where \( U_t(h_{t-1}) \) is the value of being unemployed and \( Q_t(h_t) \) is the initial value of a contract offered to a new worker in the nonfinancial sector. Although we use the same notation we used to describe the financial sector, it should be clear that the variables and value functions used here are different from those used in the description of the financial sector. For example, the variable \( \rho_t \) used here is the probability of finding a job in the nonfinancial sector. The variable \( \rho_0 \) used in the previous section, instead, is the probability of finding a job in the financial sector. Although the matching functions of the two sectors have the same form, the actual matching probabilities may differ.

To define the value of being unemployed, we assume that an unemployed worker can generate non-market consumption \( \bar{c} h_{t-1} \), where \( \bar{c} \) is exogenously given. Then the value of being unemployed is

\[ U_t(h_{t-1}) = \ln(\bar{c} h_{t-1}) + \beta E_t \left[ (1 - \rho_{t+1}) U_{t+1}(h_{t-1}) + \rho_{t+1} Q_{t+1}(h_{t-1}) \right], \quad (26) \]
where $Q_{t+1}(h_{t-1})$ is the value of an occupation in the nonfinancial sector. Since workers lose the skills to work in the financial sector if they are not employed in that sector, unemployed workers search for a job only in the nonfinancial sector. The unemployment value can be normalized to $u_t = \ln(\bar{c}) + \beta E_t[(1 - \rho_{t+1})u_{t+1} + \rho_{t+1}q_{t+1}]$.

The final object that needs to be specified is the contractual environment that prevails in the nonfinancial sector. We assume that in this sector all firms are public companies. As argued earlier, a contractual environment with public companies is better captured by double sided limited commitment. This implies that in the nonfinancial sector it is optimal for workers to chose the innovation effort that maximizes their expected outside value net of the effort cost, that is,

$$
\hat{\lambda} = \arg \max_\lambda \left\{ -e(\lambda) + \beta E_t D(h_t, \lambda, \varepsilon_{t+1}) \right\}.
$$

5.3 General equilibrium

We now have all the ingredients to define a Steady State general equilibrium. When necessary, sectorial variables and functions will be identified by superscript $j \in \{1 = \text{Finance}, 2 = \text{Nonfinance}\}$.

Definition 2 (Steady state) Given the contractual regime in the financial sector (one-sided or double-sided) and in the nonfinancial sector (double-sided), a stationary equilibrium is defined by (i) Policies $\lambda = \varphi^{\lambda,j}(q)$, $c = \varphi^{c,j}(q)$, $q(\varepsilon) = \varphi^{q,j}(q, \varepsilon)$; (ii) Normalized utilities for unemployment, $\bar{u}$, for newly hired workers, $\bar{q}$, and initial normalized value for investors, $\bar{v}^j$; (iii) Demand and price of financial services $X$ and $P$; (iv) Posted vacancies, $N^j(h)$, searching workers, $S^j(h)$, filling probabilities, $\phi^j(h)$, finding probabilities, $\rho^j(h)$; (v) Distributions of workers, $\mathcal{M}^j(h, q)$; (vi) Law of motion for the distribution of workers, $(\mathcal{M}_{t+1}^1, \mathcal{M}_{t+1}^2) = \Phi(\mathcal{M}_t^1, \mathcal{M}_t^2)$. Such that (i) The policy rules $\varphi^{\lambda,j}(q)$, $\varphi^{c,j}(q)$, $\varphi^{q,j}(q, \varepsilon)$ solve the optimal contract in each sector $j \in \{1, 2\}$; (ii) The normalized utilities $\bar{q}^j$ and investor values $\bar{v}^j$ solve the free entry condition (23) in each sector $j \in \{1, 2\}$ and investors receive a share $\eta$ of the surplus; (iii) the market for financial services clears (demand $X$ is equal to the production of financial services) and $P$ is the equilibrium price; (iv) Filling and finding probabilities satisfy $\phi^j(h) = m(N^j(h), S^j(h))/N^j(h)$ and $\rho^j(h) = m(N^j(h), S^j(h))/S^j(h)$; (v) The law of motion $\Phi(\mathcal{M}^1, \mathcal{M}^2)$ is consistent with individual policies; (vi) The distributions of workers in the financial and nonfinancial sectors are constant, that is, $(\mathcal{M}^1, \mathcal{M}^2) = \Phi(\mathcal{M}^1, \mathcal{M}^2)$.

For the later analysis, it will be convenient to state the following property:

Lemma 3 Higher values of $\rho^j$ are associated with higher steady-state contract values $\bar{q}^j$.

The lemma, proved in Appendix C, states that more competition for workers in either the financial sector or the nonfinancial sector, redistributes rents in their favor. We will use the lemma later to establish some of the main results of the paper.
5.4 Inequality

Since the income of workers depends on human capital, we can use $h$ as a proxy for the distribution of income. As a specific index of inequality we use the coefficient of variation in human capital. Since in a steady state equilibrium with double-sided limited commitment all agents choose the same $\lambda^j = \hat{\lambda}^j$, we can derive simple expressions for some of the moments of the distribution.

Denote by $\hat{\lambda}^j$ the innovation effort that maximizes the outside value of a worker in sector $j \in \{1, 2\}$. This is determined by condition (17) and it is the same for all workers in the same sector (although it is different in the two sectors). As we have seen, the equilibrium innovation in the environment with double-sided limited commitment is equal to $\hat{\lambda}^j$: since the investor cannot make credible promises, workers always choose the innovation that maximizes the outside value. Therefore, the gross growth rate of human capital for an individual worker is $g(\hat{\lambda}^j, \varepsilon)$. Appendix E shows that, in each sector $j \in \{1, 2\}$, the average human capital and the coefficient of variation for the cross sectional distribution of human capital are equal to

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

$$Std^j(h) = \sqrt{\frac{\omega[1 - (1 - \omega) Eg(\hat{\lambda}^j, \varepsilon)]^2}{\omega[1 - (1 - \omega) Eg(\hat{\lambda}^j, \varepsilon)]^2} - 1}.$$  

Therefore, the average human capital and the inequality index in each of the two sectors are simple functions of the investment $\hat{\lambda}^j$. We then have the following proposition.

**Lemma 4** With double-sided limited commitment, the average human capital and the inequality index are strictly increasing in $\hat{\lambda}^j$.

That average human capital increases with investment is obvious. The dependence of inequality on $\hat{\lambda}^j$ can be explained as follows. If $\hat{\lambda}^j = 0$, the human capital of all workers will be equal to $h_0$ and the inequality index is zero. As $\hat{\lambda}^j$ becomes positive, inequality increases for two reasons. First, since the growth rate $g(\hat{\lambda}^j, \varepsilon)$ is stochastic, human capital will differ within the same tenure cohort of workers (workers with the same employment tenure). Second, since each cohort experiences growth, the average human capital differs between cohorts of workers. 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 $\hat{\lambda}^j$. Appendix E derives analytical expressions separately for the within and between cohort inequality.

Similar properties should also hold, approximately, for the case of one-sided limited commitment. However, it is not possible to derive analytical expressions for the inequality index since the equilibrium $\lambda^j$ differs across workers of the same sector. But intuitively, the inequality index should increase with the average value of $\lambda^j$. 

25
6 The impact of organizational changes

We now explore the core issue addressed in this paper, that is, how the organizational change in the financial sector from partnerships to public companies has affected risk taking, income distribution and market valuation of firms in both sectors. We have identified two key consequences of the organizational change in the financial sector:

1. *Increased competition for human capital*: The separation between investors and managers expanded the base of potential investors who could fund a company, facilitating the expansion of businesses. In the context of our model this is captured, parsimoniously, by a reduction of the vacancy cost \( \tau \) in the financial sector. A lower \( \tau \) increases the incentive to post vacancies and, therefore, increases the competition to attract workers in the financial sector. Notice that the vacancy cost changes only in the financial sector. In the nonfinancial sector it stays the same.

2. *Weakened the commitment of investors*: While the limited commitment of managers was also a feature of the traditional partnership (partners were not prevented from leaving the partnership), the limited commitment of investors was not a problem since investors were the representatives of partners. Even from a legal standpoint, it was difficult for a partnership to replace a partner without a consensual agreement. A feature of public companies, instead, is a clearer separation between investors and managers. With this separation, the limited commitment of investors becomes also important. In the context of our model, this is captured by a shift, in the financial sector, from one-sided limited commitment to double-sided limited commitment. There are no changes, instead, in the nonfinancial sector which continues to be characterized by public companies, and therefore, double-sided limited commitment.

In summary, we formalize the demise of the traditional partnership in the financial sector as a shift to an environment with more competition for human capital and with limited enforceability also for investors. We explore first the consequences of higher competition for managers in the environment with double-sided limited commitment.

**Proposition 5** Suppose that there is double-sided limited commitment (public companies) in both sectors of the economy. Then a steady state equilibrium with a lower \( \tau \) in the financial sector (but same \( \tau \) in the nonfinancial sector) is characterized by:

1. Greater risk-taking in both sectors, that is, higher \( \hat{\lambda}^1 \) and \( \hat{\lambda}^2 \).
2. Higher share in value added and higher relative productivity of the financial sector.
3. Lower stock market valuation of financial firms, relative to nonfinancial firms.
4. Greater income inequality in both sectors.
To understand the first property we have to take into account that a lower $\tau^1$ increases the matching probability for a worker not only in the financial sector, $\rho^1$, but also in the nonfinancial sector, $\rho^2$. That a lower vacancy cost in the financial sector increases the job finding rate in this sector is obvious. This increases the outside option of workers employed in the financial sector and, with double-sided limited commitment, raises the incentive to choose a higher $\hat{\lambda}^1$.

The increase in innovation in the nonfinancial sector is less direct and follows from general equilibrium effects. As innovation in the financial sector increases, the supply of financial services raises. In the general equilibrium this generates a fall in the price of financial services, $P$, which allows for an increase in the income of nonfinancial companies. The higher profitability of nonfinancial companies increases the value of creating new jobs and this raises the job finding rate also in the nonfinancial sector. As $\rho^2$ increases, the outside value of workers also raises in the nonfinancial sector creating the incentives for choosing a higher $\hat{\lambda}^2$.

Although innovation increases in both sectors, the increase in human capital is stronger in the financial sector. This leads to an increase in the production share of the financial sector which is further amplified by the increase in the share of workers employed in the financial sector. We would like to point out that, although the increase in the share of employment would arise even if there were no contractual frictions, the increase in innovation arises only because there are contractual frictions. This is a novel feature of the 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.

The intuition for the third property—lower valuation of financial firms relative to nonfinancial firms—can be seen from the free entry condition that, in a steady state, can be written as $\phi(\bar{q})v(\bar{q}) = \tau$. If $\tau$ declines in the financial sector, the right-hand-side term also declines. In part this is obtained by a reduction in the filling probability $\phi(\bar{q})$ but in part is obtained by a reduction in $v(\bar{q})$, that is, the market valuation of a financial firm at entry. Market valuation, includes also seasoned firms, that is, firms that entered earlier. But also for these firms the value of the contract is lower because the higher probability $\rho^1$ increases the outside value of workers and, through this, they are able to extract a larger share of the surplus. In the nonfinancial sector, instead, $\tau$ does not change. Therefore, the market valuation of nonfinancial companies is not affected directly through the free entry condition. Still, the increase in the outside value of workers allow them to extract a larger share of the surplus in seasoned firms. But this effect is minor.

Finally, the fourth property—greater inequality—follows from the first property, that is, from higher values of $\hat{\lambda}^1$ and $\hat{\lambda}^2$. As we have seen in Lemma 4, a higher value of $\hat{\lambda}$ increases human capital accumulation and inequality within each sector.

The next question is how the equilibrium properties are affected by a change from one-sided limited commitment to double-sided limited commitment (keeping constant the vacancy cost $\tau$). We will characterize the effects of this change numerically since an analytical characterization is not available.
6.1 Quantitative analysis

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

The investment cost is specified as \( \kappa(\lambda) = \lambda^2 \). The production function in the nonfinancial sector is specified as \( F(x) = 1 + zx^{0.5} \). The dis-utility from innovation effort takes the form \( e(\lambda) = -\alpha \ln(1 - \lambda) \).

Given the specification of preferences and technology and after normalizing the initial human capital \( h_0 \) to 1, there are 11 parameters to calibrate (see the top section of Table 1). Given the difficulty of calibrating the parameter of the matching function \( \eta \), it is customary to set it to \( \eta = 0.5 \). We follow the same approach here (even though the matching functions are specific to the financial and nonfinancial sectors). 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.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta} ) Discount factor</td>
<td>0.962</td>
</tr>
<tr>
<td>( \alpha ) Utility parameter for dis-utility innovation effort</td>
<td>0.206</td>
</tr>
<tr>
<td>( \omega ) Death probability</td>
<td>0.025</td>
</tr>
<tr>
<td>( \psi ) Fraction newborn workers searching for jobs in finance</td>
<td>0.086</td>
</tr>
<tr>
<td>( p^1 ) Probability of successful innovation finance</td>
<td>0.036</td>
</tr>
<tr>
<td>( p^2 ) Probability of successful innovation nonfinance</td>
<td>0.063</td>
</tr>
<tr>
<td>( z ) Productivity of financial services in the nonfinancial</td>
<td>0.741</td>
</tr>
<tr>
<td>( A ) Matching efficiency</td>
<td>0.592</td>
</tr>
<tr>
<td>( \tau ) Cost of posting a vacancy in the financial sector</td>
<td>3.534</td>
</tr>
<tr>
<td>( \tau^2 ) Cost of posting a vacancy in the nonfinancial sector</td>
<td>3.534</td>
</tr>
<tr>
<td>( \eta ) Matching share parameter (pre-set)</td>
<td>0.500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calibration moments</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>0.04</td>
</tr>
<tr>
<td>Life expectancy of workers</td>
<td>40.00</td>
</tr>
<tr>
<td>Employment share in finance</td>
<td>0.06</td>
</tr>
<tr>
<td>Value added share in finance</td>
<td>0.08</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.05</td>
</tr>
<tr>
<td>Inequality index (coeff. variation) in financial sector</td>
<td>4.00</td>
</tr>
<tr>
<td>Inequality index (coeff. variation) in nonfinancial sector</td>
<td>4.00</td>
</tr>
<tr>
<td>Time allocated to innovation in finance</td>
<td>0.30</td>
</tr>
<tr>
<td>Probability of finding an occupation in finance</td>
<td>0.70</td>
</tr>
<tr>
<td>Probability of filling a vacancy</td>
<td>0.50</td>
</tr>
</tbody>
</table>

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 employment and value added shares are the approximate numbers for finance and insurance in the 2000s as shown in Figure 1. The inequality indexes comes from the 2010 Survey
of Consumer Finance for the financial and nonfinancial sectors. The last three moments (innovation time, job finding rate in finance and job filling rate) are not based on direct empirical observations and the values assigned are somewhat arbitrary. A sensitivity analysis will clarify the relevance of these calibration targets. Appendix ?? provides a detailed description of how the 10 moments are mapped into the first 10 parameters.

**Results.** Our goal is to assess the quantitative impact of greater competition and lower contract enforcement. The impact of higher competition is captured by looking at the equilibrium consequences of reducing the vacancy cost $\tau$. The impact of lower enforcement is captured by looking at the changes induced by a shift from the environment with one-sided limited commitment to the environment with double-sided limited commitment. We see the environment with one-sided limited commitment and higher vacancy cost as characterizing the financial sector in the pre-1980s period. The environment with double-sided limited commitment and lower vacancy cost is representative of recent years.

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

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

![Figure 5: Steady state investment policies for different $\tau$ in the environments with one-sided and double-sided limited commitment.](image)

In contrast, when neither managers nor investors can commit, more competition in-
duces more innovation, as Proposition 2 predicts. Also in this environment the probability of external offers increases, which raises the external value of managers and makes investment less attractive for investors. In order to implement the optimal $\lambda$, investors would need to promise adequate future compensation. The problem is that future promises are not credible with double-sided limited commitment and the only way managers can increase their contract value is by raising their outside value. This is achieved by choosing higher $\lambda$. With a lower $\tau$, the probability of an external offer $\rho$ increases. Since the manager benefits from higher innovation only if she receives an external offer, the higher probability $\rho$ raises the manager’s incentive to choose a higher value of $\lambda$.

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

<table>
<thead>
<tr>
<th>Low competition ($\tau = 5.301$)</th>
<th>One-sided limited commitment</th>
<th>Double-sided limited commitment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average value of $\lambda$ finance</td>
<td>0.186</td>
<td>0.249</td>
</tr>
<tr>
<td>Value of $\lambda$ nonfinance</td>
<td>0.166</td>
<td>0.171</td>
</tr>
<tr>
<td>Matching probability finance, $\rho$</td>
<td>0.599</td>
<td>0.627</td>
</tr>
<tr>
<td>Matching probability nonfinance, $\rho^2$</td>
<td>0.304</td>
<td>0.307</td>
</tr>
<tr>
<td>Share of employment finance</td>
<td>0.054</td>
<td>0.057</td>
</tr>
<tr>
<td>Share of output finance</td>
<td>0.068</td>
<td>0.073</td>
</tr>
<tr>
<td>Initial investor value finance, $\bar{v}$</td>
<td>9.071</td>
<td>9.501</td>
</tr>
<tr>
<td>Average investor value finance, $Ev(q)$</td>
<td>11.092</td>
<td>11.560</td>
</tr>
<tr>
<td>Coefficient of variation finance</td>
<td>0.529</td>
<td>1.026</td>
</tr>
<tr>
<td>Coefficient of variation nonfinance</td>
<td>1.461</td>
<td>1.727</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High competition ($\tau = 3.534$)</th>
<th>One-sided limited commitment</th>
<th>Double-sided limited commitment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average value of $\lambda$ finance</td>
<td>0.170</td>
<td>0.300</td>
</tr>
<tr>
<td>Value of $\lambda$ nonfinance</td>
<td>0.172</td>
<td>0.177</td>
</tr>
<tr>
<td>Matching probability finance, $\rho$</td>
<td>0.739</td>
<td>0.700</td>
</tr>
<tr>
<td>Matching probability nonfinance, $\rho^2$</td>
<td>0.307</td>
<td>0.311</td>
</tr>
<tr>
<td>Share of employment financial sector</td>
<td>0.067</td>
<td>0.063</td>
</tr>
<tr>
<td>Share of output finance</td>
<td>0.074</td>
<td>0.080</td>
</tr>
<tr>
<td>Initial investor value, $\bar{v}$</td>
<td>7.464</td>
<td>7.068</td>
</tr>
<tr>
<td>Average investor value, $Ev(q)$</td>
<td>8.643</td>
<td>8.470</td>
</tr>
<tr>
<td>Coefficient of variation finance</td>
<td>0.476</td>
<td>4.000</td>
</tr>
<tr>
<td>Coefficient of variation nonfinance</td>
<td>1.765</td>
<td>2.400</td>
</tr>
</tbody>
</table>

So far we have shown that the organizational change that took place in the financial sector induced more risk-taking. We now show that they also generated other changes that are consistent with the observations we highlighted in the introduction. Table 2 shows that the shift to an environment with double-sided limited commitment and lower $\tau$ is associated with some increase in the share of employment in the financial sector and a bigger increase in the share of output. Another important prediction of the model is that the shift is associated with a reduction in the (average) value of investors, relative to human capital. Since we do not have physical capital, we use human capital as a proxy.
for the book value of assets.\footnote{This would be the case if we explicitly introduce capital and assume that there is complementarity between human and physical capital.} Table 2 also shows that the initial investor’s value is lower. This follows directly from Lemma 3 and the free entry condition \( \phi(q) \cdot v(q) = \tau \) after the reduction in the vacancy cost \( \tau \).

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

Finally, we emphasize that, even if there are no structural changes in the nonfinancial sector, innovation and income inequality increase also in this sector. This is because the increase in the supply of financial services reduces the price of these services. As the price declines, more services are used in the production of final goods, which increases the income of the nonfinancial firms. As nonfinancial income raises, the incentive to create jobs increases, increasing the probability of a match also in the nonfinancial sector (higher competition for human talents). It is the increase in the matching probability in the nonfinancial sector that induces higher values of \( \lambda \) in this sector.

7 Conclusion

The financial crisis of 2007-2009 has brought attention to the growth in size and importance of the financial sector over the past few decades, as well as the increase in risk taking in the financial sector. 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 argue that changes in the organizational structure of financial firms have increased competition for managerial skills and weakened the enforcement of contractual relationships between workers/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 and the weakened commitment 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. The focus on the financial sector is
motivated by three considerations. First, the organizational changes described in the
introduction have been more evident in the financial sector. Second, we also believe that
some of the features of this sector—that our model helps to explain—are less present in
other sectors (for example the relatively low book value). Third, managerial talent
is an extremely important factor of production in the financial sector and it is particularly
inalienable (capital and unskilled labor play a more relevant role in other innovative
sectors and patents on financial instrument are rare avis and difficult to enforce).9

However, the modeling of the financial sector is general and can be used to study
similar organizational changes in other sectors. The general predictions of the model
is that, when organizations are subject to external competition—with different effects
on members of the organization—competition is likely to distort internal decisions and
result in redistribution of ex-post rents. With enough commitment (in our model, one-
sided limited commitment), the organization can internalize these distortions but this
does not mean it can implement the ex-ante full-commitment allocation which makes the
organization immune to ex-post competition (with one-sided limited commitment there
is lower risk-taking in response to competition).

It can be argued that modern financial organizations have many credible instruments
(bonuses, etc.) to overcome the investor’s commitment problem and, therefore, that our
model with two-sided limited commitment is a poor description of innovative financial
firms. But we have explicitly chosen to work with a simplified model in order to sharpen
the key mechanism that emerges in the presence of limited commitment. Sophisticated
compensation packages for CEOs and workers in general are just partial forms of limited
commitment compared to the internal compensation schemes that dominated in the
previous organizational form, that is, the traditional partnership.10

9 Although these differences with other innovative sectors may be a question of degree “But perhaps the most significant change has been to human capital. Recent changes in the nature of organizations, the extent and requirements of markets, and the availability of financing have made specialized human capital much more important, and also much more mobile. But human capital is inalienable, and power over it has to be obtained through mechanisms other than ownership”. Rajan and Zingales (2000).
10 “The highest incomes and the largest fortunes in the financial sector were made by investing one’s money—in other words, as a partner of a private bank rather than as a manager of a joint stock bank.” Cassis (2013).
Appendices

A Proof of Proposition 1

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

$$\min_{\chi, \gamma(\varepsilon')} \max_{c, \lambda} \left\{ \beta \pi(\lambda) - c + \mu \left( u(ch) - e(\lambda) \right) - \chi \left( e(\lambda) - e(\hat{\lambda}) \right) \right. $$
$$\left. + \beta E \left[ v(\mu') g(\lambda, \varepsilon') + \left( \mu + \chi + \gamma(\varepsilon') \right) Q(h', \mu') \right] \right. $$
$$\left. - \chi D \left( h, g(h', \varepsilon') h, \mu' \right) - \gamma(\varepsilon') D(h, h', \rho) \right\}$$

s.t. $h' = g(\lambda, \varepsilon') h$, $\mu' = (\mu + \chi + \gamma(\varepsilon')) / g(\lambda, \varepsilon')$,

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

$$(\mu + \chi) e_\lambda(\lambda) - \beta \pi_\lambda(\lambda) \geq \beta E \left[ \left( v(\mu') + (\mu + \chi + \gamma(\varepsilon')) Q_h(h', \mu') \right) \right. $$
$$\left. - \gamma(\varepsilon') D_2(h, h', \rho) \right]$$

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

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

$$u(C) = \frac{C^{1-\sigma}}{1-\sigma}$$
the optimal consumption policy, (10), takes the form: 
\[(ch)^{-\sigma} = (h\mu)^{-1}\]
that is,
\[u(h\mu) = \frac{(h\mu)^{\frac{1-\sigma}{\sigma}}}{1-\sigma},\]
and, therefore,
\[u_{h,\mu}(h\mu) = \frac{1-\sigma}{\sigma^2} (h\mu)^{\frac{1-2\sigma}{\sigma}}.\]
In sum, \[u_{h,\mu}(h\mu) \leq 0\] if and only if \[\sigma \geq 1\]; i.e. if and only if the intertemporal elasticity of substitution is less or equal one. Otherwise, if \[1/\sigma > 1\] the optimal contract will tend to lower current consumption in exchange for compensating the manager in the future with \[Q_{h,\mu}(h\mu) > 0\]. Notice that, given the separability between consumption and effort \[Q\] inherits its differentiability properties from \[u\] (we abstract from some technicalities in making this claim). We analyse in detail the particular case of \(\sigma = 1\); i.e. \(Q_{h,\mu}(h\mu) = 0\).

B The first-order conditions of Problem (18)
Let \(\mu, \gamma(\varepsilon)\) and \(\chi\) be the lagrange multipliers associated with the promise-keeping, enforcement and incentive-compatibility constraints. The lagrangian can be written as
\[v(q) = \beta(1 - \lambda) - c + \beta \sum_{\varepsilon} g(\lambda, \varepsilon) v(q(\varepsilon)) p(\varepsilon)\]
\[+ \mu \left\{ \ln(c) - e(\lambda) + \beta \sum_{\varepsilon} \left[ B \ln \left( g(\lambda, \varepsilon) \right) + q(\varepsilon) \right] p(\varepsilon) - q \right\}\]
\[+ \beta \sum_{\varepsilon} \left[ q(\varepsilon) + (1 - \rho) B \ln \left( g(\lambda, \varepsilon) \right) - \bar{d} \right] \gamma(\varepsilon) p(\varepsilon)\]
\[+ \chi \left\{ -e(\lambda) + \beta \sum_{\varepsilon} \left[ B \ln \left( g(\lambda, \varepsilon) \right) + q(\varepsilon) \right] p(\varepsilon) - d(\hat{\lambda}) \right\}.\]

The terms \(\bar{d}\) and \(d(\hat{\lambda})\) collect variables and functions that are not affected by the contract policies \(\lambda, c\) and \(q(\varepsilon)\). The first order conditions with respect to these three variables are, respectively,
\[-\beta - (\mu + \chi) e_1(\lambda) + \beta \sum_{\varepsilon} g_1(\lambda, \varepsilon) v(q(\varepsilon)) + B \left( \frac{g_1(\lambda, \varepsilon)}{g(\lambda, \varepsilon)} \right) \left( \mu + \chi + (1 - \rho) \gamma(\varepsilon) \right) p(\varepsilon) = 0\]
\[-1 + \frac{\mu}{c} = 0\]
\[g(\lambda, \varepsilon) v_1(q(\varepsilon)) + \left( \mu + \chi + \gamma(\varepsilon) \right) = 0\]
Substituting the envelope condition $v_1(q) = -\mu$ and using the functional forms of $\pi(\lambda)$ and $g(\lambda, \varepsilon)$ we obtain equations (20)-(21).

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

$$\rho_t(h, Q_t(h)) \cdot \left[ Q_t(h) - Q_t(h) \right] = \rho_t(h, Q^*_t(h)) \cdot \left[ Q^*_t(h) - Q_t(h) \right],$$

(29)

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

$$\max_{Q_t(h)} \left\{ \phi_t(h, Q_t(h)) \cdot V_t(h, Q_t(h)) \right\}$$

subject to (29),

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\(^{11}\) $V_t(h, Q_t(h))$.

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

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

$$\phi_t(h, Q_t(h)) \cdot V_t(h, Q_t(h)) = \tau h.$$

(31)

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 $h$.

\(^{11}\)Given the free entry condition, the ‘initial value’ for the investor is 0 and the initial value of the contract is, in fact, his ‘interim value’, but when there is no confusion we also refer to the initial value of the contract as the ‘initial value’.
that is, \( V_t(h, Q_t(h)) = v_t(q_t)h_t \). Therefore, the free entry condition can be rewritten in normalized form as

\[
\phi_t(\bar{q}_t) \cdot v_t(\bar{q}_t) = \tau.
\] (32)

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 \( \phi_t \))\(^{12}\).

The investor’s problem (30) can be rewritten as

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

subject to

\[
\rho_t(q)(q - q_t) = \rho_t(\bar{q}_t^*)(\bar{q}_t^* - q_t), \quad \forall q \in N(\bar{q}_t^*)
\]

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

\[
1 - \eta = \frac{-v'_t(\bar{q}_t)(\bar{q}_t - q_t)}{v_t(\bar{q}_t) - v'_t(\bar{q}_t)(\bar{q}_t - q_t)},
\] (33)

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 \( v_e'(\bar{q}) \) denote the elasticity of the investor’s value function; i.e. \( v_e'(\bar{q}) \equiv -\frac{v'(\bar{q})\bar{q}}{v(\bar{q})} \). Our log-linear specification implies that \( v_e'(\bar{q}) > 0 \).

**Lemma 3A** \( v_e'(\bar{q}) > 0 \) implies \( \bar{q}'(\rho) > 0 \).

The optimality condition (33) can be written as

\[
\frac{1 - \eta}{\eta} = v_t(q) \frac{\bar{q} - q}{\bar{q}}.
\] (34)

In a stationary equilibrium, using (33) we obtain:

\(^{12}\)In equilibrium only skilled workers who have never been employed in the financial sector will be actively searching. Since they have never been employed in the financial sector, they all have human capital \( h_0 \). For determining the probability of a match when a financial manager decides to quit, we incur the problem that the number of posted vacancies is discrete. In this case we assume that investors randomize over the posting of a vacancy that is targeted at a manager with human capital \( h \).
\[
\bar{q} - q = \bar{q} - \{ \ln(1) + \beta \left[ (1 - \rho)q + \rho \bar{q} \right] \}
\]
\[
= (1 - \beta) \bar{q} + \beta (1 - \rho) (\bar{q} - q)
\]
\[
= (1 - \beta (1 - \rho))^{-1} (1 - \beta) \bar{q};
\]

therefore

\[
v_{\epsilon}(\bar{q}) = \frac{1 - \eta}{\eta} \frac{\bar{q}}{\bar{q} - q}
\]
\[
= \frac{1 - \eta}{\eta} \frac{(1 - \beta (1 - \rho))}{(1 - \beta)}.
\]

Taking derivatives with respect to \( \rho \),

\[
v_{\epsilon}'(\bar{q}) \bar{q}'(\rho) = \frac{1 - \eta}{\eta} \frac{\beta}{1 - \beta} > 0;
\]

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

D The numerical solution

We describe first the numerical procedure used to solve Problem (18) for exogenous outside values \( q \) and \( \bar{q} \) and for exogenous probability of offers \( \rho \). 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

\[
\mu = \psi(q) \]
\[
v = \Psi(q).
\]

The first function returns the multiplier \( \gamma \) (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

$$
\beta \left[ v(q(1)) + \left( \frac{B}{1 + \lambda} \right) \left( \mu + \chi + (1 - \rho)\gamma(1) \right) \right] p = -\beta \pi\lambda(\lambda) + \frac{\alpha(\gamma + \chi)}{1 - \lambda} \tag{35}
$$

$$
c = \gamma \tag{36}
$$

$$
g(\lambda, \varepsilon)\psi(q(\varepsilon)) = \mu + \chi + \gamma(\varepsilon) \tag{37}
$$

$$
v = \beta \pi(\lambda) - c + \beta \sum_{\varepsilon} g(\lambda, \varepsilon)\Psi(q(\varepsilon))p(\varepsilon) \tag{38}
$$

$$
q = \ln(c) + \alpha \ln(1 - \lambda) + \beta \sum_{\varepsilon} \left( B \ln \left( g(\lambda, \varepsilon) \right) + q(\varepsilon) \right) p(\varepsilon) \tag{39}
$$

$$
\chi \left\{ \alpha \ln(1 - \lambda) + \beta \sum_{\varepsilon} \left[ q(\varepsilon) + B \ln \left( g(\lambda, \varepsilon) \right) \right] p(\varepsilon) \right. 
- \left. \alpha \ln(1 - \hat{\lambda}) - \beta \sum_{\varepsilon} \left[ (1 - \rho)\hat{q} + \rho\hat{q} + \rho B \ln \left( g(\hat{\lambda}, \varepsilon) \right) \right] p(\varepsilon) \right\} = 0 \tag{40}
$$

$$
\gamma(\varepsilon) \left[ q(\varepsilon) - (1 - \rho)\hat{q} - \rho\hat{q} + (1 - \rho)B \ln \left( g(\lambda, \varepsilon) \right) \right] = 0 \tag{41}
$$

The first three equations are the first order conditions with respect to $\lambda$, $c$, $q(\varepsilon)$, respectively. Equation (38) defines the value for the investor and equation (39) is the promise-keeping constraint. Equations (40) and (41) formalize the Kuhn-Tucker conditions for the incentive-compatibility and enforcement constraints.

Notice that equations (40) and (41) must be satisfied for all values of $\varepsilon$ which can take two values. Therefore, we have a system of 9 equations in 9 unknowns: $\lambda$, $c$, $v$, $\mu$, $\chi$, $q(\varepsilon)$, $\gamma(\varepsilon)$. Once we have solved for the unknowns we can update the functions $\psi(q)$ and $\Psi(q)$ using the solutions for $v$ and $\mu$.

**Solving for the steady state equilibrium.** The iteration starts by guessing the steady state values of $\hat{q}$ and $\rho$. Given these two values, we can determine $q$ using equation (35). 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 $\hat{q}$ and $\rho$ using the free-entry condition (32) and the bargaining condition (33). We keep iterating until convergence, that is, the guessed values of $\hat{q}$ and $\rho$ are equal to the computed values (up to a small approximation error).
E Derivation of the inequality index

With double-sided limited commitment in both sectors, the inequality index takes the same form. Therefore, we derive the index without specifying the sector.

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

\[
\bar{h} = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_k h_k, 
\]

where \( E_k \) averages the human capital of all agents in the \( k \)-cohort. Because growth rates are serially independent, we have that \( E_k h_k = h_0 E g(\hat{\lambda}, \varepsilon)^j \). 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

\[
\text{Var}(h) = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_k (h_k - \bar{h})^2.
\]

This can be rewritten as

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

Using the serial independence of the growth rates, we have that \( E_k h_k^2 = h_0^2 [E g(\hat{\lambda}, \varepsilon)^2]^k \). 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 standard deviation (square root of (43)) by the mean (defined in (42)). This returns the inequality index (27).

We can separate the within and between components of the inequality index starting by rewriting the formula for the variance of \( h \) as,

\[
\text{Var}(h) = \omega \sum_{k=0}^{\infty} (1-\omega)^k \left[ (E_k h_k^2 - \bar{h}^2) - (h_k^2 - \bar{h}^2) \right],
\]

39
where \( \bar{h}_k = E_k h_k = h_0 E g(\hat{\lambda}, \varepsilon)^k \) is the average human capital for the \( k \) cohort. Substituting the expression for \( h_k \) and \( \bar{h}_k \) 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 (42), we are able to write the square of the inequality index as

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

(44)

The first term is the \textit{within} cohorts inequality while the second term is the \textit{between} cohorts inequality. Both terms are strictly increasing in \( \hat{\lambda} \).
References


Bell, Brian and John Van Reenen. 2010. “Banker’s Pay and Extreme Wage Inequality in the UK,” Centre for Economic Performance, LSE.


