Endogenous agency problems and the dynamics of rents

Bruno Biais (TSE) and Augustin Landier (HEC Paris)[1]

July 2018

Abstract

While potentially more productive, more complex tasks request more effort, generating larger agency rents. Agents therefore prefer to acquire complex skills, to earn large rents. In our overlapping generations model, their ability to do so is kept in check by competition with predecessors. Old agents, however, are imperfect substitutes for young ones, because the latter are easier to incentivize, thanks to longer horizons. This reduces competition between generations, enabling young managers to go for larger complexity than their predecessors. Consequently, equilibrium complexity and rents gradually increase, especially when agents are patient and turnover limited, so that compensation deferral is very useful to mitigate moral hazard.

Keywords: Agency rents, moral hazard, dynamic contracts, complexity.

JEL codes: D8, G2, G3.

[1]This research was initiated within the Paul Woolley Research Initiative on Capital Market Dysfunctionalities at IDEI, Toulouse School of Economics. Many thanks for insightful feedback to the editor, Dimitri Vayanos, and three anonymous referees as well as participants at seminars in Toulouse, Lausanne, Lugano, University of Virginia, NYU Abu Dhabi, City University, MIT Sloan, the 4^{th} Game Theory Society world meeting in Istanbul, the NBER, the 6^{th} Paul Woolley Conference at the LSE, and the FIRS conference in Dubrovnik, especially Philip Bond, Matthieu Bouvard, Max Bruche, Bruce Carlin, Thomas Mariotti, Roger Myerson, Guillaume Plantin, Larry Samuelson, Lucio Sarno, David Sraer and Jean Tirole. Bruno Biais benefitted from the support of the European Research Council, grant 295484 TAP.

1 Introduction

Agency problems arise when principals cannot precisely observe and control what agents do (Holmstrom, 1979; Grossman and Hart 1983; Holmstrom and Tirole, 1997). To provide incentives to limited liability agents, principals must promise large compensation in case of success. This gives rise to agency rents.[2]

As shown in Holmstrom and Tirole (1997), agency rents increase with the cost of effort, or, equivalently, the private benefit from shirking. The (opportunity) cost of effort varies with the characteristics of the task delegated to the agent. An important characteristic of a task is its complexity. Brünnermeier and Oehmke (2009) note that a complex problem can be decomposed in sequence of simple, elementary ones. To successfully complete the complex task, the agent must solve each of the subproblems. Hence, the larger the complexity of the tasks and the number of subproblems, the larger the agent's effort, and the more severe the moral hazard problem. They analyse variation in the number of tasks a manager has to carry out. Dessein and Santos (2006) emphasize the flexibility enjoyed by managers with broad job definitions. While Dessein and Santos (2006) do not consider agency problems, this suggests that complex jobs, with many tasks and large flexibity, leaving significant discretion to the agent, are exposed to more severe moral hazard. Combining these observations with Holmstrom and Tirole (1997), complex tasks raises more severe moral hazard problem, and generate larger agency rent.

We study the dynamics of complexity and rents in an equilibrium model in which successive generations of agents acquire increasingly sophisticated skills corresponding to increasingly complex tasks, generating increasing rents.

Model and results: To clarify the origin of rents in our analysis, we assume there is no scarcity of managers. Thus, if the market for managers was frictionless, principals would hire only those managers that are optimal from their point of view, maximizing returns net of rents. Since agents

[2] Instead of unobservable ex-ante effort, Thomas and Worall (1988), Kocherlakota (1998), Townsend (1979), Diamond (1984), and Bolton and Scharfstein (1990), emphasize ex-post unobservability and limited commitment: A problem arises when the agent cannot commit not to leave the firm, while it would be difficult to complete the task successfully if the agent absconded. In this context, the principal must leave the agent a rent, to convince him not to abscond. Our key insights also obtain in that alternative, but for our purpose essentially equivalent, framework.

would choose to acquire only those skills that make them employable, complexity would not rise above what is optimal for principals. In contrast, we assume there are search frictions in the labour market.

We consider an overlapping generations model, in which agents live two periods. At the beginning of his life, a generation t young manager chooses (at a cost) a given skill, corresponding to a given type of task, denoted by b. b can be interpreted as the number of layers of complexity of the task. More complex tasks potentially generate larger gross returns, but their completion also demands larger efforts, and therefore entails larger costs.

Once agents have acquired their skill, each young principal meets a young agent, observes his b, and decides whether to hire him or not.[3] When making this choice, the principal bears in mind that she could instead *i)* search for another generation t agent or *ii)* hire a generation $t-1$ agent, and then another agent at time $t+1$. We assume principals incur a (possibly very small) cost when searching for managers. This shuts down competition between contemporaneous managers, as in Diamond (1971), enabling one to focus on the key driving force in our model: competition between successive generations.

It is particularly attractive for a generation t principal to try and hire a generation $t-1$ agent if low bs were chosen by that generation. Thus, when generation $t-1$ acquired skills corresponding to simple tasks, this limits the ability of generation t to increase its own b to earn high rents. The competitive pressure imposed by the previous generation is limited, however, by the *endogenously* imperfect substitutability among generations. The intuition is the following: To reduce rents, principals defer compensation (as in Becker and Stigler, 1974, and Rogerson, 1985). This makes it relatively unattractive for a young principal to hire an old agent. Indeed, old agents have short horizons, which prevents deferring their compensation to reduce their rents. Thus, other things equal, it is cheaper to incentivize young agents than old ones. Because old agents are imperfect substitutes for young ones, the latter can afford to choose technologies with greater agency problems than their predecessors, and still be hired. This gives rise to an upward trend in complexity and agency rents, which is the core result of our paper, stated in Proposition 1, below. Note that, even when higher complexity raises *gross* returns, the increase in rents eventually reduces *net* returns for principals.

[3] In our main analysis, we focus on the case in which the principal can make a take-it or leave-it offer to the agent. We then relax this assumption, analysing the case in which the agent has some, but not all the, bargaining power. Our key results still hold in that extension.

Empirical implications: The main novel empirical implication of our theoretical analysis is that increasing complexity should be associated with increasing agency rents. This should be particularly pronounced for industries in which agents have greater opportunities to opt for complex techniques, and in which such complexity is more likely to create agency problems.

One industry for which this is particularly relevant is finance, where the lack of hard–wired technological constraints raises the scope for rent-seeking driven complex innovations.[4] Accordingly, Philippon and Resheff (2008) observe a simultaneous increase in managers' rents and investment techniques complexity, while Célerier and Vallée (2014) document an increase in complexity for structured products, and Greenwood and Sharfstein (2012) observe an increasing share of institutions relying on complex investment techniques. Also in line with our model, Böhhm, Metzger and Strömberg (2018) observe rising rents in the finance sector without an increase in talent and suggest this points to moral hazard rents.

In order to confront our model to a broader cross-section of industries, we use S&P Capital IQ Professional Data, which document job functions and compensation for professionals. We focus on US firms between 2010 and 2016, a period during which the capital IQ dataset is well documented. We rely on two proxies for complexity: the number of functions of an executive and the number of occurrences of the term "complexity" in a company's 10-K form. For both proxies, we find that, during our sample period, average compensation grew significantly more in industries with larger increase in complexity.[5] We also find, in line with our model, that rents grew more in industries with lower turnover.[6]

Our model delivers additional testable predictions: The increase in rents and complexity spurred by an initial deregulation or technology shock should not be instantaneous. Rather it should be sustained and delayed. Consequently sustained increases in rents and complexity can still take place when there is no current change in exogenous variables.

Yet another implication of our analysis relates the increase in rents to the search for yields. In

[4] Another industry in which increasingly complex techniques may have led to increasing rents is the healthcare industry (see e.g. Bodenheimer, 2005).

[5] We do not claim causality. Our empirical results should be seen as descriptive and illustrative.

[6] Turnover acts in our model as a restriction on long-term contracts, which mitigates the ability of agents to extract more rents than previous generations.

general, the rise of complexity and rents is limited by the constraint that agents must leave enough return to the principals to convince them to delegate the management of their wealth rather than self-invest it. When the return on self-investment is low, which can be proxied by low safe return and low return on indexing, this should increase agents' ability of agents to increase complexity and rents.

Finally, our analysis implies that experienced managers and junior managers are imperfect substitutes. This should show up in hiring and compensation data. For example, when new slots open up, experienced managers are imperfect substitutes for junior ones. Our theory also predicts that imperfect substitutability, and its consequences, should be stronger when incentive problems are more severe and compensation more backloaded.

Literature: Our analysis of agents' rent–seeking is in line with Baumol (1990) and Murphy, Shleifer and Vishny (1991). Both in their analysis and ours, rent–seeking agents impose costs upon the others. In Baumol (1990) and Murphy, Shleifer and Vishny (1991), however, these costs are exogenously directly induced by the actions of the rent–seeker, e.g., warfare, litigation or predatory trading. In contrast, in our analysis, the initial choice of the agent (complexity) has an indirect endogenous impact on the principal, via the agency rent it induces, and also on subsequent increases in complexity.[7]

Our work is also related to Axelson and Bond (2015)'s equilibrium analysis of dynamic contracting with overlapping generations and moral hazard. In Axelson and Bond (2015) agents can be assigned to two types of task, with different levels of moral hazard and productivity. Thus, a common theme in their paper and ours is the selection of tasks, and corresponding moral hazard, arising in equilibrium. The endogenous rise in rents over time, reflecting imperfect competition between successive generation, is one of the key specific results of our model, differentiating it from Axelson and Bond (2015).

Our point that agents in the finance industry choose complex products and techniques to increase the rents they extract from principals, echoes the point made by Carlin (2009) that competing financial institutions design complex products to increase their market power. A major difference is that our analysis hinges on agency problems, which can arise even with large rational investors, while Carlin (2009) focuses on retail investors and abstracts from agency issues.

[7]In Rajan and Zingales (1998), agents' ex-ante investment in human capital increases their rents ex-post. While in Rajan and Zingales (1998) this can be efficient, in our analysis there is inefficiently large investment in technology.

Our analysis is also related to Bolton, Santos and Scheinkman (2013). In their paper also, opportunistic occupational decisions lead to rents and inefficiencies and give rise to externalities. The economic mechanisms at work in the two papers are different, however. In Bolton et al (2013), that many agents choose to become dealers in the OTC market worsens adverse selection in the other market. In contrast, in our analysis, that agents choose complex techniques worsens moral hazard, and increases rents, for the following generations.

Last, our paper is related to the literature on social norms, which explores, notably, how parents transmit values or preferences to their children (see, e.g., Bisin and Verdier, 2000). In our model, in contrast with that literature, the transmission of norms from one generation to the next is driven by competition between generations. And we show that the imperfection of that competitive process induces a decline in standards.

The next section presents the model. Section 3 presents the optimal contract designed by one principal, hiring one agent for two periods. Section 4 embeds this bilateral contracting problem in an equilibrium labour market context and analyzes the dynamics of rents. Section 5 discusses robustness and extensions. Section 6 presents our empirical analysis. Section 7 briefly concludes. Proofs not given in the text are in the appendix.

2 Model

Investors and managers: Each period, a mass–one continuum of investors and a mass–M continuum of managers are born. $M \geq 1$, so that there is no scarcity of managers. In this overlapping generations model, successive generations of managers coexist in the market at a given point in time, which creates the scope for competition between generations.[8]

All market participants are risk neutral, have limited liability and live for two periods. The discount factor of investors is $\rho \in (0, 1)$, while that of managers is $\beta \in (0, 1)$. In line with the literature on dynamic financial contracting (e.g., DeMarzo and Fishman (2007) and Biais, Mariotti, Plantin and

[8] While successive generations of agents are key to our analysis, our qualitative results would be unchanged if we considered infinitely lived principals.

Rochet (2007)), we assume $\rho \geq \beta$.

Each investor is initially endowed with one unit of investment good. She can invest it in a default technology, which she can operate herself and which returns 1 unit of consumption good per period during two periods. Alternatively, she can delegate the management of her capital to an agent, hereafter referred to as "the manager". For simplicity the choice between self–investment and delegated investment is irreversible.

Managers have zero initial endowment. At the beginning of his life, each young manager must choose among a range of techniques indexed by $b \in [0,1]$. Each technique corresponds to a specific type of skills, knowhow and human capital. The (non–monetary) cost of acquiring skills b is equal to cb, with $c \geq 0$. Importantly, the choice of b is irreversible. The idea is that managers acquire skills, human capital, relations and technical knowledge at an early stage in their career. Then, they use this informational capital.[9]

Complex tasks: When entrusted with one unit of capital, manager b can generate return equal to $R(b) \geq 1$ units of consumption good per period during each of the two periods of his life. We assume R is continuous, increasing and concave in b.[10] $R'(b)$ denotes the left derivative of $R(b)$. Since R is concave, R' is decreasing and, in the same spirit as Inada conditions, we assume that $R'(1) = 0$.

b measures the complexity of the task delegated to the agent. As noted by Brünnermeier and Oehmke (2009, page 6): "One way to deal with complexity is by dividing up a larger, complex task into smaller, more manageable subtasks." b can be thought of as the number of subtasks. The greater the number of subtasks to be completed, the greater the cost of effort for the agent.

As Holmstrom and Tirole (1997), we assume effort is unobservable and the agent has limited liability, which raises a moral hazard problem. At the beginning of each period, the agent can exert effort or shirk. When the agent exerts effort, i.e., checks each of the b layers of complexity, the project generates cash flow $R(b)$ for sure, while, when the agent fails to exert effort, cash flows can be $R(b)$ with probability $1 - \Delta$, or 0 with probability Δ.

In line with Holmstrom and Tirole (1997), instead of framing the model in terms of cost of effort,

[9] See Oyer (2008) for empirical evidence on long term effects of initial career paths in the financial sector.

[10] The assumption that R is increasing is not needed for our analysis. Our qualitative results are upheld when R is constant. Moreover, one can interpret $R(b)$ as the efficient frontier of the production set, in b/output plane.

we, equivalently, assume private benefits from shirking, i.e., opportunity cost of effort.[11] In line with the above discussion, we assume this opportunity cost of effort is increasing in b, the complexity of the task. More precisely, we assume the private benefit is equal to $b\Delta R(b)$, i.e., fraction $b \in [0,1]$ of the loss in expected output due to shirking ($\Delta R(b)$). Increasing complexity makes checking all the aspects of the project more demanding, which raises the opportunity cost of effort.[12]

Sequence of play: for an investor and a manager born at $t \geq 1$, the timing of actions is the following:

At time t:

- Stage 1: Young manager i in generation t chooses $b_t^i \in [0,1)$.
- Stage 2: Each young investor is matched with one young manager, observes his b_t^i, and decides whether to make him a take–it–or–leave–it contract offer or reject him. For simplicity, we assume the principal has all the bargaining power. In Subsection 5.2.1 we show that our results are robust to giving the agent some (but not all) bargaining power. Since there is a mass one of investors, and a mass $M \geq 1$ of managers, each manager is matched with an investor with probability $1/M$. This probability is the same for all managers. In particular, it cannot depend on the choices made by managers at stage 1, because an investor can check a manager's b only after being matched with him.
- Stage 3: If the investor decides to reject the young manager with whom she was matched, or if the manager rejects the offer, then the investor decides whether to self–invest or search for another manager, at cost ϵ, which can be arbitrarily small. If the investor decides to continue searching for managers, she can direct her search towards young or old managers. Then, on meeting a new manager, the investor observes his b and can make him a take–it–or–leave–it offer, and the process is iterated. Eventually, investment takes place.
- Stage 4: Each employed manager decides whether to exert effort or not, then output is realized (and equals $R(b_t^i)$ or 0), and the manager receives the compensation stated in the contract.

[11] The difference is that, while the cost of effort is incurred on the equilibrium path, the private benefits from shirking are not. This leads to slightly simpler expressions.

[12] Sato (2015) offers a more precise microfounded analysis of the effect of complexity and opacity on agency problems.

At time $t+1$:

- If the investor born at time t hired a manager born at time t, they apply contract signed at time t. If the time t output was 0 and the contract stated the manager should be fired in that case, the investor can search for a new agent, at cost ϵ.
- If the investor born at time t hired a manager born at time $t-1$, she can search for a new manager at $t=1$.
- If the investor did not hire a manager at time t, she invested her endowment of investment good in the default technology, and therefore no longer has any choice to make at $t+1$.

3 Optimal contracting

In this section, we analyse the optimal contract designed by one principal, hiring one agent, taking b_t as given. In the next section we embed this contracting problem in a market equilibrium and study the endogenous determination of b_t. The compensation contract offered at time t by the investor states the wages to be received by the manager as a function of the output realized at time t and at time $t+1$. It also specifies if the manager should be kept after period t or fired.

First consider the contract requesting the agent to exert effort at both times. In that case, on the equilibrium path, the output is equal to $R(b_t)$ at each period. It is clearly optimal to fire the manager, without any compensation, when output is 0. Hence, the time t contract is pinned down by the pair of wages, w_t^t and w_{t+1}^t, paid to the manager if output $R(b_t)$ is generated in period t and in period $t+1$. After success at time t, the incentive compatibility condition at time $t+1$ is that the gain of the agent when exerting effort (wage w_{t+1}) be larger than or equal to his gain when shirking (wage with probability $1-\Delta$ plus private benefit from shirking)

$$w_{t+1}^t \geq (1-\Delta)w_{t+1}^t + \Delta b_t R(b_t),$$

that is

$$w_{t+1}^t \geq b_t R(b_t). \tag{1}$$

At the end of period t, after $R(b_t)$ has been obtained, the continuation utility of the agent, anticipating effort at $t+1$, is βw_{t+1}. Thus, at time t the incentive compatibility condition is

$$w_t + \beta w_{t+1}^t \geq (1-\Delta)(w_t^t + \beta w_{t+1}^t) + \Delta b_t R(b_t).$$

That is

$$w_t^t + \beta w_{t+1}^t \geq b_t R(b_t). \tag{2}$$

In this section we set the exogenous outside option of the manager to 0, so that his participation constraint never binds. In the next section, the endogenous outside option of the manager will still be 0. The program of the investor is to maximize expected net returns subject to incentive compatility constraints, i.e.,

$$\max_{w_t^t, w_{t+1}^t} R(b_t)(1+\rho) - w_t^t - \rho w_{t+1}^t, \text{ s.t., (1) and (2).} \tag{3}$$

The solution to this program is spelled out in the next lemma.

Lemma 1: *At time t, for a given choice of b_t, if $\rho > \beta$, the solution to (3) is such that (1) and (2) bind, the wages conditional on success are*

$$\{w_t^t, w_{t+1}^t\} = \{(1-\beta)b_t R(b_t), b_t R(b_t)\}. \tag{4}$$

If $\rho = \beta$, only (2) binds, but the wages in (4) are still (weakly) optimal. In the optimal contract inducing effort the net gains of the investor are

$$Z(b_t) \equiv R(b_t)[(1+\rho) - (1+\rho-\beta)b_t], \tag{5}$$

while the present value of the manager's earnings is

$$w_t^t + \beta w_{t+1}^t = b_t R(b_t). \tag{6}$$

Complexity, net returns and rents. By Lemma 1, the present value of the fund manager's earnings is $b_t R(b_t)$. Thus, the agent captures a fraction (b_t) of the gross return on the investment over one period ($R(b_t)$). Since the agent's outside reservation utility is 0, $b_t R(b_t)$ is his rent. Although the principal has all the bargaining power at the contracting stage, incentive compatibility and limited liability imply the agent earns a rent. In that sense, the agency problem (parametrized by b) gives the agent some endogenous bargaining power.

As the complexity of the task (b_t) increases, the total gross return increases, because $R' \geq 0$. In addition, the fraction of that return captured by the agent also increases, because the agency problem worsens. Combining these two effects, the compensation of the agent increases with the complexity of the task.

While agents benefit from an increase in complexity, principals can be made better off or worse off when complexity rises. Indeed,

$$Z'(b_t) = [(1+\rho) - (1+\rho-\beta)b_t]R'(b_t) - (1+\rho-\beta)R(b_t).$$

Because $(1+\rho) > (1+\rho-\beta)b_t$ and $R^{'}(b)$ is decreasing, the investors' net return is concave in the complexity of the task. And because $R'(1) = 0$, we have $Z'(1) \leq 0$. Thus, starting from $Z(0)$, investors' net return initially increases with b, reflecting the increase in gross return $R(b)$. Then, it reaches a maximum point at

$$b^* = \arg\max_b Z(b_t). \tag{7}$$

Finally, for $b > b^*$, investors' net return goes down with b, reflecting that an increasing fraction of the return is captured by the agent. To make things interesting, we assume $Z'(0) \geq c$, i.e., at the lowest level of complexity the enhancement in net return brought about by an increase in b exceeds the cost c. Finally denote by $b_{\max}$ the highest value of b in $[0,1]$ such that $Z(b) \geq 1+\rho$, i.e., investors prefer delegated investment rather than self-investment.

Optimality of effort: For the contract in Lemma 1 to be the optimal contract, it must generate higher net gains for the principal than the alternative contracts requesting i) no effort at all, or ii) effort at time t and no effort at time $t+1$, or iii) no effort at time t and effort at time $t+1$. By Lemma 1, the net gains of the investor requesting effort at both periods are as given in (5). On the other hand, if the principal lets the agent shirk at both periods, he does not need to pay any wage, and his net gains are

$$R(b_t)(1+\rho)(1-\Delta). \tag{8}$$

(5) is greater than (8) iff

$$\Delta \frac{1+\rho}{1+\rho-\beta} \geq b_t. \tag{9}$$

Since $b_t \in [0,1]$, this always holds iff Δ is large enough, in the sense that

$$\Delta \geq \frac{1+\rho-\beta}{1+\rho}. \tag{10}$$

One can show that under (10) shirking once is also dominated by effort at both periods. Hence, the contract spelled out in Lemma 1 is the optimal contract if (10) holds, which, for simplicity, we assume hereafter. If we did not make that assumption, there would be a threshold value of b at which the principal would prefer to give up on effort. In equilibrium, agents would not set b above that threshold, whose role would be similar to that of $b_{\max}$. Apart from that, relaxing (10) would not alter our results.

Example: A simple example is when $R(b)$ is the piecewise linear function $\min[\alpha b + 1, \bar{R}]$, where α and $\bar{R} > 1$ are positive constants. For this simple case, if the agency problem is not too severe, in the sense that

$$\frac{1+\rho}{1+\rho-\beta} \geq \frac{2\bar{R}-1}{\alpha}, \tag{11}$$

then the level of complexity maximising $Z(b)$ is also that maximising output net of cost,

$$b^* = \frac{\bar{R}-1}{\alpha}. \tag{12}$$

4 Equilibrium dynamics

We now turn to the dynamics of complexity. We focus on symmetric equilibria, in which all managers born at time t choose the same equilibrium level of complexity, b_t^*.

For simplicity, we assume the ex-ante expected gain of an agent, $\frac{bR(b)}{M} - bc$, is increasing in b, i.e.,

$$\frac{R(b) + bR'(b)}{M} \geq c. \tag{13}$$

This implies that, for any b, the ex-ante expected gain of an agent is non negative, and also that any $b < b^*$ would be Pareto dominated: both managers and investors would be better off with a larger b. So, we initialize the process at $b_0 = b^*$.

From that point on, any increase in complexity reduces the net returns of investors, while raising the rents of managers. Thus, there is a conflict of interest between the former and the latter. We now study whether market forces lead to an equilibrium that is more favorable for the investors (keeping complexity at b^*) or for the managers (letting complexity rise above b^*).

Given an initial level of complexity, b_0, an equilibrium is a sequence $\mathcal{E} = \{b_t^*, w_t^{t^*}, w_{t+1}^{t^*}\}_{t\geq 1}$, satisfying the following conditions:

- **Optimization:** At each time t, each young manager i chooses b_t^i to maximize his gains, and each investor makes an optimal hiring decision.

- **Rational expectations:** Investors and managers have rational expectations about the equilibrium dynamics $\mathcal{E}$ and find it optimal to also play according to $\mathcal{E}$. Thus, on the equilibrium path at time t, young manager i finds it optimal to set $b_t^i = b_t^*$, and each investor offers the optimal contract

$$\{w_t^{t*}, w_{t+1}^{t*}\} = \{(1-\beta)b_t^* R(b_t^*), b_t^* R(b_t^*)\}. \tag{14}$$

In each generation, at stage 2, each manager is drawn with probability $\frac{1}{M}$. This probability does not vary with managers' bs, because we assume that, before contacting the manager, the investor cannot observe the manager's type. Once drawn, a manager strictly prefers to be hired and earn (6). Thus, at stage 1, manager i chooses b_t^i to maximize his expected gains

$$\frac{1}{M} R(b_t^i) b_t^i - c b_t^i, \tag{15}$$

subject to the constraint that the investor prefers hiring hime when drawing him. To analyse that constraint, we need to compare the investor's payoff when hiring the young manager to her payoff from alternative actions:

- The first alternative option for the investor is self–investment. She does not choose that option if her net return on delegated investment, $Z(b_t^i)$, is larger than

$$1 + \rho. \tag{16}$$

- The second alternative option for the investor is to hire an old agent in period t and then hire a generation t manager at $t+1$. At time t she would have to compensate the old agent enough to avoid shirking. This would entail promising the old agent compensation at least as large as $b^*_{t-1}R(b^*_{t-1})$. Such compensation would attract the old manager irrespective of whether he is employed or not. Similarly, at time $t+1$ the investor would have to pay the new recruit $b^*_t R(b^*_t)$. Hence, overall, if she were to opt for that deviation, the time t investor would expect to get

$$R(b^*_{t-1})(1-b^*_{t-1})+\rho R(b^*_t)(1-b^*_t)-\epsilon(1+\rho), \tag{17}$$

where the last term ($\epsilon(1+\rho)$) is the search cost of going after an old manager at t and then another one at $t+1$.

- The third alternative option for the investor is to hire an old manager at t and then a young one at $t+1$. In this case, when deviating, the generation t investor expects to pay $b^*_{t-1}R(b^*_{t-1})$ to the old manager she hires at time t, and $b^*_{t+1}R(b^*_{t+1})$ to the young manager she hires at time $t+1$. Consequently, the deviating investor expects to earn

$$R(b^*_{t-1})(1-b^*_{t-1})+\rho R(b^*_{t+1})(1-b^*_{t+1})-\epsilon(1+\rho). \tag{18}$$

- The fourth alternative option for the investor is to search for another young manager at time t, expecting to hire him for two periods and to compensate him with $\{w^{t*}_t, w^{t*}_{t+1}\}$ given in (14). In this case the investor expects to earn

$$Z(b^*_t)-\epsilon. \tag{19}$$

Overall, the employability constraint for the young manager is that $Z(b^i_t)$ be larger than or equal to (16), (17), (18), and (19).

Bearing in mind that $b_{\max}$ is the highest b such that investors prefer delegated investment rather than self-investment, the young manager is employable as long as he picks b in the subset of $[b^*, b_{\max}]$ such that

$$Z(b)\geq \max[R(b^*_{t-1})(1-b^*_{t-1})+\rho\max[R(b^*_t)(1-b^*_t),R(b^*_{t+1})(1-b^*_{t+1})]-\epsilon(1+\rho),Z(b^*_t)-\epsilon].$$

Since the expected gain of the young manager is increasing in b as long as he remains employable, his optimal choice is as stated in the next lemma.

Lemma 2: *The maximisation program of the young agent at time t has a unique solution b_t which is either equal to $b_{\max}$ or such that*

$$Z(b_t) = \max[R(b^*_{t-1})(1 - b^*_{t-1}) + \rho \max[R(b^*_t)(1 - b^*_t), R(b^*_{t+1})(1 - b^*_{t+1})] - \epsilon(1 + \rho), Z(b^*_t) - \epsilon]. \quad (20)$$

Equation (20) implicitly defines the function ϕ giving the optimal choice of b_t as a function of b^*_{t-1}, b^*_t and b^*_{t+1}, i.e., $b_t = \phi(b^*_{t-1}, b^*_t, b^*_{t+1})$. In equilibrium, the young investor must find it optimal to choose a level of complexity equal to b^*_t. Therefore, either $b^*_t = b_{\max}$ or

$$b^*_t = \phi(b^*_{t-1}, b^*_t, b^*_{t+1}). \quad (21)$$

It can never be the case that $Z(b^*_t) = Z(b^*_t) - \epsilon$, even when ϵ is arbitrarily small, as long as it is strictly positive. Thus the cost of searching for managers, even if it is very small, shuts down competition within the same generation.[13] Consequently, evaluated at $b_t = b^*_t$, (20) simplifies to

$$Z(b^*_t) = R(b^*_{t-1})(1 - b^*_{t-1}) + \rho \max[R(b^*_t)(1 - b^*_t), R(b^*_{t+1})(1 - b^*_{t+1})] - \epsilon(1 + \rho). \quad (22)$$

$R(b)(1 - b)$ is the net return to an investor hiring a manager, with skill b, for one period. For $b \geq b^*$, this net return is decreasing with b, reflecting that the manager extracts an increasing fraction of the surplus.[14] Thus, when it is expected that complexity will increase from t to $t + 1$, the max in (22) is $R(b^*_t)(1 - b^*_t)$. In that case, either $b^*_t = b_{\max}$ (in which case $b^*_t \geq b^*_{t-1}$), or (22) simplifies to

$$R(b^*_t)(1 - b^*_t) - R(b^*_{t-1})(1 - b^*_{t-1}) = -\beta b^*_t R(b^*_t) - \epsilon(1 + \rho). \quad (23)$$

Since, the right–hand–side of (23) is negative, we have that $R(b^*_t)(1 - b^*_t) \leq R(b^*_{t-1})(1 - b^*_{t-1})$, that is $b^*_t \geq b^*_{t-1}$. Thus, between $t - 1$ and t, there is an increase in complexity, worsening agency problems, and eroding investors' returns while raising managers' rents. This is stated in the next lemma.

[13] As discussed below, this is similar to Diamond (1971), but, in contrast with Diamond (1971), in our model, managers from generation t also compete with their predecessors and successors.

[14] To see this note that $Z'(b)$, which for $b \geq b^*$ is negative, is equal to the derivative of $R(b)(1 - \delta b)$ plus a positive term.

Lemma 3: *If $b^*_{t+1} \geq b^*_t$, then $b^*_t \geq b^*_{t-1}$.*

To interpret the increase stated in Lemma 3, consider the right–hand–side of (23). $\epsilon(1+\rho)$, the cost incurred by investors searching for another manager, obviously limits competition between managers, in particular managers belonging to the same generation (as in Diamond, 1971). To focus on the specific economic mechanism at play in our model, which is driven by competition between managers from different generations, consider the limit case where ϵ goes to 0. In that case, the increase in b is solely driven by $\beta b^*_t R(b^*_t)$. This is the difference between the net investor's revenue when the principal hires agent b^*_t on a long term basis $(Z(b^*_t))$ and when she hires the agent via a sequence of short–term contracts $((1+\rho)R(b^*_t)(1-b^*_t))$. This is a measure of the advantage of long term contracting, which is feasible with young agents, but not with old ones. Thus, it is a measure of the extent to which old managers are only imperfect substitutes for young ones. (23) shows how young managers take advantage of this imperfect substitutability: They raise complexity (and thus rents) above the prior level, up to the point at which investors are indifferent between i) hiring young managers on a long–term basis to complete a more complex task, and ii) hiring old managers on a short–term basis to complete a less complex task.

While Lemma 3 spells out what happens at time t when b is expected to rise after t, the next lemma states that future bs cannot decrease in equilibrium.

Lemma 4: *b^*_t never decreases.*

The intuition for Lemma 4 is the following. By Lemma 3, if b was to decrease at t, it would have to decrease at $t+1$. In fact, as shown in the appendix, b would have to go down by increasingly large lumps and eventually go below 0, which is a contradiction. Combining Lemmas 2, 3 and 4, we obtain our first proposition:

Proposition 1: *There exists a unique symmetric equilibrium. In that equilibrium, investors hire managers from their own generation for two periods. Equilibrium complexity and agents' rents increase until b^*_t reaches $b_{\max}$. Starting from $b_0 = b^*$, as long as $b^*_t < b_{\max}$, b^*_t is the unique solution of the recursive equation (23), which implicitly defines the function ψ mapping b^*_{t-1} into b^*_t.*

Proposition 1 directly implies the next corollary, which gives a lower bound on the increase in b_t^* due to imperfect competition among generations.

Corollary 1: *As long as $b_t^* < b_{\max}$, the growth of b_t^* is faster than exponential, i.e.,starting from $b_0 = b^*$,*

$$b_t^* \geq \frac{b^*}{(1-\beta)^t}. \tag{24}$$

The greater the patience of the agent (β), the greater the advantage of long–term contracts over short term–contracts, the lower the substitutability among generations, the higher above b_{t-1}^* generation t can raise b_t^*. Hence the larger the lower bound on the growth of b_t^*.

Relation with Diamond (1971) and role of ϵ. If ϵ was strictly equal to 0, the equilibrium condition would be

$$Z(b_t^*) = \max[1+\rho, R(b_{t-1}^*)(1-b_{t-1}^*) + \rho\max[R(b_t^*)(1-b_t^*), R(b_{t+1}^*)(1-b_{t+1}^*)], Z(b_t^*)]. \tag{25}$$

The process b_t^* presented in Proposition 1 solves (25), and thus remains an equilibrium when $\epsilon = 0$. There are, however, other equilibria, in which the equilibrium value of b is between b^* (defined in (7)) and b_t^* (characterized in Proposition 1). In those equilibria, it is the last term (rather than the middle one) that binds in the max on the right–hand–side of (25). That is, the choice of b by a generation t manager is constrained by the choices of his competitors from the same generation (not by those of his predecessors).

When $\epsilon = 0$ and (25) holds, it is weakly optimal for investors not to resample after drawing a manager. If they follow that strategy, the equilibrium remains as in Proposition 1. It is, however, also weakly optimal for an investor to sample all the $1 - M$ managers that are not employed, after drawing a manager with b_t^*. If a manager anticipated such behaviour, then his best–response would be to opt for b_t^i slightly lower than b_t^*, to make sure he would eventually be drawn and hired. Since all managers would reason similarly, this would drive the equilibrium choice of b down to b^*. In this type of equilibrium, competition between managers would lead to the outcome preferred by investors.

This argument, however, and the possibility for b^* to be an equilibrium, don't apply when ϵ is strictly positive, in which case the unique equilibrium is that characterized in Proposition 1. Thus

ϵ, arbitrarily close to, but strictly above, 0, limits competition between managers belonging to the same generation. This is comparable to the way search costs limit competition and generate rents in Diamond (1971). One contribution of our analysis, relative to Diamond (1971), is to study the equilibrium dynamics of rents, and show they have a tendency to increase along the equilibrium path. To see this more clearly, note that the model in Diamond (1971) is similar to a one–period version of our model, where the equilibrium condition on the level of b prevailing at time 1 would be

$$Z(b_1^*) = \max[1 + \rho, Z(b_1^*) - \epsilon]. \tag{26}$$

(26) immediately leads to $b_1^* = b_{\max}$. This contrasts with our model where b_t^* progressively increases over several periods, before eventually reaching $b_{\max}$. The reason why the increase in b_t^* is only progressive in our model is that the choice of generation t is constrained by the choices of previous generations. That anchor does not exist in Diamond (1971). Yet, in our model, the moderating effect of the previous generation is limited, due to imperfect substitutability between generations. Hence the gradual increase in b_t^*.

Externalities. When choosing b_{t-1}^*, generation $t-1$ sets a benchmark, with which generation t will have to compete when choosing b_t^*. Thus, while the actions of generation $t-1$ have no *exogenous direct* effect on the following generation, they exert an *endogenous externality* on the latter. When choosing a relatively high level of complexity b_{t-1}^*, generation $t-1$ does not internalize that this will lead to an even larger level of complexity b_t^*, and thus large rents for generation t managers and low net returns for generation t investors.

Compensation and seniority. Lemma 1 implies that $w_t^t < w_{t+1}^t$. Thus, for a given generation, compensation rises with seniority, i.e., a given agent earns more when senior than when junior. Proposition 1 and Corollary 1, however, imply that senior managers from the previous generations earn less than junior managers from the current generation. Indeed, from Lemma 1, $w_t^{t-1*} \leq w_t^{t*}$ iff

$$b_{t-1}^* R(b_{t-1}^*) \leq (1-\beta) b_t^* R(b_t^*). \tag{27}$$

Since $b_t^* \geq b_{t-1}^*$, (24) implies (27). The increase in rents (driven by the increase in complexity) from one generation to the next is larger than the increase in compensation, within one generation, from one period to the next.

Example: In our simple example, in which $R(b) = \min[\alpha b + 1, \bar{R}]$, the following corollary obtains:

Corollary 2: *If $R(b) = \min[\alpha b + 1, \bar{R}]$ and (11) holds then, as ϵ goes to 0, b_t^* goes to*

$$\min[\frac{b^*}{(1-\beta)^t}, b_{\max}]. \tag{28}$$

In general, the increase in b_t^* above b^*, made possible by the imperfect substitutability between old and young managers, is enhanced by the fact that $R(b)$ increases in b. In the simple example, however, $R(b)$ is constant when b is above b^*. In that situation, the increase in b_t^* is solely due to the imperfect substitutability between old and young managers. Correspondingly, the growth in b_t^* is just equal to its lower bound, stated in Corollary 1.

Welfare. As discussed above, the optimal level of complexity for investors is b^*, which maximizes $Z(b)$ (see (7)), while (by (13)) it is $b_{\max} \geq b^*$ for managers. Now turn to what a benevolent social planner would decide. For simplicity, we herafter set $\rho = \beta$. Since utilities are linear, there is a unique Pareto optimum regarding real decisions, and the points on the Pareto frontier differ only in terms of purely redistributive transfers between investors and managers. In the first best, the social planner solves the following problem:

$$\max_{b \in [0,1]} W(b) = (1+\beta)R(b) - Mcb.$$

The optimum is such that the marginal benefit of effort equals its marginal cost, i.e.,

$$b^{**} = R'^{-1}(\frac{Mc}{1+\beta}).$$

Since $R^{'}(1) = 0$, we have $b^{**} \leq 1$. Now,

$$W(b) = Z(b) + (bR(b) - Mcb).$$

Hence

$$\frac{\partial W(b)}{\partial b}|_{b=b^*} = (R(b^*) - Mc) + b^* R'(b^*),$$

which, by (13), is positive. Consequently, $b^{**} \geq b^*$. We summarize this discussion is the next proposition:

Proposition 2: *The level of complexity preferred by investors is lower than the socially optimal level of complexity, which, in turn, is lower than the level of complexity preferred by the managers, i.e.,*

$$b^* \leq b^{**} \leq b_{\max}.$$

The fraction of total surplus obtained by investors is decreasing in b. Therefore they prefer b to be lower than the social optimum. In contrast, the fraction of total surplus obtained by managers is increasing in b, and therefore they want it to be higher than the social optimum.

Delayed adjustment to changes in the environment. Suppose complexity increased, according to the law of motion given in Proposition 1, and reached its maximum level: $b_{\max}$. For simplicity, we hereafter focus on the simple case in which $R(b)$ increases linearly with slope α until $\bar{R}$ and then becomes flat for all $b \geq b^* = \frac{\bar{R}-1}{\alpha}$. This implies $b_{\max} = (1+\beta)(1-\frac{1}{\bar{R}})$.

Now assume that, due to a technological breakthrough or change in regulation, for projects initiated from time t on, $\bar{R}$ is raised to $\frac{\bar{R}}{\omega}$, with $\omega \in (0,1)$.[15] This leads to an increase in the maximum possible level of complexity to

$$(1+\beta)(1-\frac{\omega}{\bar{R}}).$$

How does this affect complexity and rents?

The next proposition states that the positive shock generates a progressive increase in complexity and rents, which persists until all the additional profitability has been absorbed by agency rents.

Proposition 3: *Suppose that, at time $s < t$, equilibrium complexity reached $b_{\max} = (1+\beta)(1-\frac{1}{\bar{R}})$ and then remained constant. If, at time t, there is a one-off permanent technological change, raising $\bar{R}$ up to $\frac{\bar{R}}{\omega}$, then, if $\beta\bar{R} \leq 1$, complexity starts rising again until it reaches its new maximum $(1+\beta)(1-\frac{\omega}{\bar{R}})$.*

At $t-1$, the constraint that the principal be better off hiring an agent than self investing was binding. The time t technological shock relaxes this constraint. Yet, complexity does not jump immediately to the new level at which the constraint binds. Rather, the positive technological shock

[15] Our analysis is valid irrespective of whether this change was anticipated or not.

implies that now it's the employability constraint that binds complexity choices. Thus b_t^* is anchored by b_{t-1}^*, and progressively rises, as in Proposition 1, until it reaches its new maximum. Our theory thus predicts that increases in rents and complexity will occur long after (and not just after) technology or deregulation shocks.

5 Robustness and extensions

In this section we discuss extensions of the model and compare it to an alternative specification. For simplicity we hereafter focus on the case, introduced above as an example, in which $R(b)$ is the piecewise linear function $\min[\alpha b + 1, \bar{R}]$, where α and $\bar{R} > 1$ are positive constants.

5.1 Agents' bargaining power

In the analysis above we assumed the principal had all the bargaining power. This implied that an agent's rent, when hired at time t, was set by the binding incentive compatibility constraint, and therefore just equal to $b_t^* \bar{R}$. How would relaxing this assumption affect equilibrium dynamics?

First, consider the case in which agents have *all* the bargaining power. Then, when a principal contacts an agent, the latter demands all the surplus, leaving the former with the reservation utility from self–investment, $1 + \beta$. If the principal was to reject such a demand, then she would have to draw another agent, at cost ϵ. Since this new agent would also have all the bargaining power, he would also offer the principal $1 + \beta$ only. Hence, rejecting the initial offer and drawing another agent would give the principal at most $1 + \beta - \epsilon$. Thus, when agents have all the bargaining power and there are (possibly infinitesimal) search costs, principals only obtain $1 + \beta$, while agents extract any additional value creation (in line with the Diamond paradox). In this context, rents immediately jump to their maximum level: In contrast with the benchmark case in which principals have all the bargaining power, there is no progressive increase in rents. Moreover, while in the benchmark case, agents strategically opt for complexity in order to obtain rents, in the alternative case in which agents have all the bargaining power, agents choose complexity to maximise productive efficiency.

Second, consider the case in which agents have some of the bargaining power, but not all. More precisely, assume that, each time a principal and an agent meet, with probability $\phi < 1$ the agent can make a take it or leave it offer and, with probability $1 - \phi$, the principal can make a take it or leave it offer (in our benchmark case in the previous section, $\phi = 0$.)

When meeting at time t a young agent i with complexity b_t^i, if the principal has the bargaining power, she can offer her $b_t^i \bar{R}$, since, if the young agent refused, he would get 0 (as he would be rejected and never drawn again.) In contrast, at the initial meeting time, if the young agent has all the bargaining power, she demands rent w_t^{ϕ}, leaving the principal with

$$\bar{R}(1 + \beta) - w_t^{\phi}.$$

If the principal rejects that offer, one thing she can do is to draw another young agent, hoping to have the bargaining power with that one. If she does have the bargaining power she gets

$$\bar{R}(1 + \beta) - b_t^* R.$$

If she does not, the agent offers her $w_t^{\phi *}$, and if she accepts, she gets:

$$\bar{R}(1 + \beta) - w_t^{\phi *},$$

where $w_t^{\phi *}$ is the equilibrium wage the principal expects to pay if he draws another agent. Hence, to be acceptable, the initial offer of the young agent must be such that

$$\bar{R}(1 + \beta) - w_t^{\phi} \geq \phi(\bar{R}(1 + \beta) - w_t^{\phi *}) + (1 - \phi)(\bar{R}(1 + \beta) - b_t^* \bar{R}) - \epsilon,$$

where the left-hand side is what the principal gets if she accepts the young agent's offer, and the right-hand side is a lower bound on what the principal expects to get if she rejects the initial agent and draws another one from the pool.

In equilibrium, $w_t^{\phi *} = w_t^{\phi}$. Thus the acceptability constraint yields

$$b_t^* \bar{R} + \frac{\epsilon}{1 - \phi} \geq w_t^{\phi *}.$$

Since, at the same time, incentive compatibility requires

$$w_t^{\phi *} \geq b_t^* \bar{R},$$

we have

$$b_t^* \bar{R} + \frac{\epsilon}{1 - \phi} \geq w_t^{\phi *} \geq b_t^* \bar{R}.$$

Thus, as long as ϕ remains bounded away from 1 (i.e., agents don't have all the bargaining power), when ϵ goes to 0, the agency rent goes to $b_t^* \bar{R}$. Because the principal can draw another agent at infinitesimal cost ϵ, even if the agent has some bargaining power, she obtains just $b_t^* \bar{R}$, exactly as when the principal has all the bargaining power. Hence, the results obtained in the previous sections, corresponding to $\phi = 0$, still hold when $1 > \phi > 0$: There is a progressive increase in rents, driven by the equilibrium increase in complexity.

5.2 Experience and productivity

In our benchmark model, young agents are more attractive than old ones, because they have longer horizon. This effect could be undermined if experience increased productivity, which would increase the attractivity of older agents relative to younger ones. To examine this point we now assume that, when exerting effort, a young agent generates $R(b)$, while an old one generates $\theta R(b)$ with $\theta > 1$. While experience increases productivity, ageing in itself does not. This rules out an equilibrium in which only old agents would be employed, since these agents would not be experienced and therefore would not be productive. Thus, in equilibrium young agents are employed for two periods, as in the benchmark case.

In this context, first take $b \geq b^*$ as given and consider a principal dealing with an agent over two periods. It is easy to show that the optimal contract is to pay the agent only when output $\bar{R}$ is obtained, and set wages equal to

$$w_t = b \max[1 - \beta\theta, 0]\bar{R},$$

in the first period, and

$$w_{t+1} = b\theta\bar{R},$$

in the second period. Thus, when old agents are really more productive than young ones, as $\theta > \frac{1}{\beta}$, the agent is paid only at the second period, and, at the first period, the present value of his rent is $\beta b\theta\bar{R}$. In contrast, when $\theta \leq \frac{1}{\beta}$, the agent is paid at both periods and, at the first period, the present value of his rent is $(1 - \beta\theta)b\bar{R} + \beta b\theta\bar{R} = b\bar{R}$, as in our benchmark model.

First consider the case in which the productivity advantage of experience is moderate, i.e., $1 < \theta \leq \frac{1}{\beta}$. Following the same logic as in our benchmark model, the employability condition for the generation

t agent choosing his b_t is (as ϵ goes to 0)

$$\bar{R}(1+\beta\theta) - b_t\bar{R} \leq \bar{R}(\theta - b^*_{t-1}\theta) + \beta\bar{R}\max[\theta - b^*_t\theta, 1 - b^*_{t+1}].$$

The left–hand side is the present value of the gains of the principal hiring for two periods the young agent he drew from the pool. The right–hand side is the present value of the gains of the principal rejecting the young agent and then hiring two consecutive agents. The first term on the right-hand side is the net gain of the principal at the first period, during which she hires an old agent from generation $t-1$. The second term on the right–hand side is the present value of the gains of the principal at the second period. The max operator reflects that the principal will optimize between hiring an old agent (from generation t) and a young agent (from generation $t = 1$).

Binding the employability condition, and imposing the equilibrium condition that $b_t = b^*_t$, we have

$$\bar{R}(1+\beta\theta) - b^*_t\bar{R} = \bar{R}(\theta - b^*_{t-1}\theta) + \beta\bar{R}\max[\theta - b^*_t\theta, 1 - b^*_{t+1}].$$

Simplifying, this yields

$$b^*_t - b^*_{t-1} = \frac{(\theta - 1 + \beta\theta)b^*_{t-1} - (\theta - 1)}{1 - \beta\theta} + \frac{\beta}{1-\beta\theta}\min[0, (\theta - 1) + b^*_{t+1} - b^*_t\theta].$$

If

$$b^*_{t-1} > \frac{\theta - 1}{\theta - 1 + \beta\theta},$$

then the right-hand side is strictly positive, and therefore $b^*_t > b^*_{t-1}$. Thus we can state our next proposition:

Proposition 5: *If $\beta\theta \leq 1$ and*

$$b^* > \frac{\theta - 1}{\theta - 1 + \beta\theta},$$

then equilibrium complexity increases until it reaches $b_{\max}$.

So when the productivity advantage of experienced agent is moderate, the equilibrium outcome is qualitatively similar to what obtained in the benchmark model. Now, turn to the alternative case, in which $\theta > \frac{1}{\beta}$. In that case, the present value of the rent of a young agent b hired for two periods is $\beta b\theta\bar{R}$. Thus, the employability condition is

$$\bar{R}(1+\beta\theta) - \beta b_t\theta\bar{R} \leq \bar{R}(\theta - b^*_{t-1}\theta) + \beta\bar{R}\max[\theta - b^*_t\theta, 1 - \beta\theta b^*_{t+1}].$$

In equilibrium this must hold for $b_t = b^*_t$. This yields the equilibrium condition

$$\bar{R}(1+\beta\theta) - \beta b^*_t\theta\bar{R} \leq \bar{R}(\theta - b^*_{t-1}\theta) + \beta\bar{R}\max[\theta - b^*_t\theta, 1 - \beta\theta b^*_{t+1}].$$

Simplifying

$$b^*_{t-1}\theta \leq (\theta - 1) + \beta\max[0, \theta(b^*_t - \beta b^*_{t+1}) - (\theta - 1)].$$

Since b^*_t cancels out, except in the max on the right-hand side, the employability constraint does not impose a cap on b^*_t. Hence, the only constraint on that choice is that the principal prefer delegated management rather than self investment. Correspondingly, b^*_t moves directly to $b_{\max}$.

Thus, the equilibrium prevailing when θ is large is qualitatively different from the benchmark case. Because experienced agent are much more productive than rookies, there is no competition between generations, and the maximum value of complexity and rents is reached immediately.

5.3 Turnover

Now suppose that, at the end of period t, with probability $1-\lambda$, an agent born at the beginning of t is hit by an exogenous shock forcing him/her to leave the market. The principal who initially hired this agents must, at period $t+1$ hire a new agent. To do so the principal draws from the pool of unemployed old agents, offering wage $b^*_t R(b^*_t)$. For the agents born at t and who have ot been hit by a shock, the incentive compatibility condition at $t+1$ is still (1). In contrast, at time t, the incentive compatibility condition is no longer (2), but

$$w^t_t + \lambda\beta w^t_{t+1} \geq b_t R(b_t). \quad (29)$$

As before, the two incentive conditions bind and the wage profile of the agent born at time t is

$$\{w^t_t, w^t_{t+1}\} = \{(1-\lambda\beta)b_t R(b_t), b_t R(b_t)\}, \quad (30)$$

while the net return to the principal is

$$R(b_t)[(1+\rho) - (1+\rho-\lambda\beta)b_t]. \tag{31}$$

Thus, everything is as in the benchmark model, except that the discount factor β is now multiplied by the probability to be around at the next period, λ. Thus, as ϵ goes to 0, b_t^* goes to

$$\min[\frac{b^*}{(1-\lambda\beta)^t}, b_{\max}], \tag{32}$$

and, noting that the turnover rate is $1-\lambda$, we can state the next corollary:

Corollary 3: *When an agent can be hit by an exogenous shock forcing him/her to leave the market, the equilibrium growth rate of rents is decreasing in the probability of this shock, which, in equilibrium, is the turnover rate.*

5.4 An alternative model with technological progress but without moral hazard

The analysis in the previous sections shows that moral hazard combined with endogenous choice of complexity leads to an increase in rents and complexity. Could the same patterns obtain, without moral hazard, just because of technological progress? To examine that issue, consider an alternative model differing from ours in three ways:

First, there is technological progress. To capture this as simply as possible, assume $R(b) = \min[\alpha b_t + 1, \bar{R}_t]$ where $\bar{R}_t$ increases with time.

Second, there is no moral hazard.

Third, to ensure that the agent can earn rents (in spite of the absence of moral hazard), assume wages are pinned down by Nash bargaining: The fraction of gross profit ($R(b_t)$) transferred as wage to the agent is equal to the bargaining power of the agent, which we denote by γ.

In this context, the principal and the agent agree on the optimal level of complexity: that which maximises output net of costs. Since $\alpha > c$, the optimum is obtained at $b_t^* = \frac{\bar{R}_t - 1}{\alpha}$. To compare the equilibrium outcomes in that alternative model and in ours, it is useful to focus on the following four variables:

- Complexity: In both models, complexity increases with time: b_t^* is $\frac{b^*}{(1-\lambda\beta)^t}$ in our model and $\frac{\bar{R}_t-1}{\alpha}$ in the alternative model.

- Wages: In both models, agent's total wage increases with time: $b_t^* R(b_t^*)$ in our model and $\gamma R(b_t^*)$ in the alternative model.

- Wages as a fraction of gross output: In our model, agent's total wage as a fraction of gross output is b_t^*, which increases with time, while in the alternative model it is equal to γ, a constant.

- Deferral: In our model, to mitigate moral hazard agent's pay must be partly deferred, so that $w_t^t < w_{t+1}^t$, in spite of the fact that the agent is more impatient than the principal. In the alternative model, because the agent is more impatient than the principal, wages are front loaded, so that $w_t^t > w_{t+1}^t$.

6 Empirical analysis

Our model delivers two types of implications about the relation between managers' compensation on the one hand, and complexity and turnover on the other hand:

First, *at a bilateral contracting level* our model predicts that individuals whose jobs are more complex, and hence more prone to moral hazard, should have higher compensation, reflecting agency rents (see Lemma 1).

Second, *at an equilibrium level* our model predicts that industries in which complexity grew more should also be industries in which compensation grew more (see Proposition 1). Our model also predicts that this increase in compensation should be muted in industries in which turnover is large (see Corollary 3).

A formal econometric test of our model is beyond the scope of this paper. Rather, the goal of this section is to present empirical patterns in compensation, complexity and turnover and compare them to the above described implications from theory.

6.1 Data

S&P Capital IQ Professional Data: Our main data source is S&P Capital IQ Professional Data (accessed through WRDS), which gathers profiles of professionals with their company affiliations. Available data includes individual identifiers, standardized job functions, titles and compensation. Using information on the start date and end date of the affiliation on an individual with a given company, we build a panel where each observation is an (individual, firm, year) triplet. We restrict the data set to employees of publicly-traded firms that were included at least once, between 2010 and 2016, in the S&P Composite 1500 Index.[16] We focus on the recent period 2010-2016 because Capital IQ is substantially weaker before 2010.

In our sample, in a given year, a firm employs on average 20 professionals included in the Capital IQ dataset.[17] We define a yearly measure of turnover at the firm level by computing for each (firm, year) the number of relationships ending during the year divided by the total number of observable employees.

Proxies for complexity: Our first proxy for the complexity of an individual's job is the number of functions of that individual in Capital IQ. The larger the number of functions of the agent, the larger the number of layers of complexity, in the same spirit as in Brünnermeier and Oehmke (2009). Also the larger the number of functions, the larger the flexibility and discretion enjoyed by the manager, in the same spirit as in Dessein and Santos (2006). Thus, the larger the number of functions, the more severe the moral hazard.

Our second proxy for complexity is computed at the firm-level, by counting the number of occurrences of the string "complex" in a firm's 10-K report: When firm complexity is higher, we expect the word "complexity" to be reported more frequently. And we expect managers operating in complex firms to have complex tasks, raising the scope for moral hazard. To construct this second proxy for complexity, we access and parse company filings information from the U.S. Securities and Exchange

[16] This index combines three indices, the S&P 500, the S&P MidCap 400, and the S&P SmallCap 600 to cover approximately 90% of the U.S. market capitalization.

[17] Thus the number of employees that we can attach to a company at a given point in time is substantially higher than in EXECUCOMP. This is one of the reasons why we use these data. The second one is that Capital IQ reports the Number of Functions of each employee, which, as discussed next, we use to proxy for complexity.

Commission (SEC) available online in the SEC EDGAR system. For each company in our sample, we collect the 10-K filing in 2010 and in 2016, when available, and calculate the number of occurrence of the string "complex" in the document.

Our two proxies of complexity are quite different: The first one varies with individuals, so that it offers some variation within firms. The industry-level correlation of the two proxies is .12, showing that, while positively correlated, the two variables are capturing different dimensions of complexity.

Variables used in the empirical analysis: In the appendix, we provide a detailed definition of the all variables used in the empirical analysis. The median total annual compensation of the professionals in our sample is USD 765,113, with a standard deviation of USD 1,403,558. The median number of functions is 3, with a standard deviation of 2.13. The median number of occurrences of the term "complexity" in a firm's 10-K is 3, with a standard deviation of 6.37. The median yearly turnover rate is 10%, with a standard deviation of 4%.

We compute variables at the SIC 3 industry level by aggregating information of firms-workers in each industry, and averaging across firms in the industry. Thus we compute average total compensation growth, average variation in the number of functions of workers and average change in the number of occurrence of "complex" in the 10-K filings, between 2010 and 2016, as well as average yearly turnover.

Furthermore, to be used as control variables, we calculate the average market value of firms in a given industry, as well as corresponding average market value growth between 2010 and 2016, using Compustat consolidated company-level market value (common shares outstanding multiplied by the month-end price that corresponds to the period end date).

6.2 Regressions

Individual-level regressions: In line with Lemma 1, we regress the log of an individual's average annual compensation over the years spent at a given firm on our two complexity variables. For the first proxy of complexity (NbFunction), which varies within firms, we can add firm fixed effects, which allows for an identification based on within firms variations in job complexity. In line with the model, and for both proxies, there is a significantly positive link between complexity and the level of total compensation.[18] To give a sense of the economic magnitude, a one standard deviation change in the second measure of complexity (NbComplex10K) is associated with a 13% increase in the level of total compensation.

Industry-level regressions: In a second set of regressions, reported in Table 2, we compare aggregate behaviours across industries. To do so, we average variables at the industry level (SIC3). When computing averages, we restrict ourselves to individuals working in the same SIC3 industry in 2010 and 2016. Last, we only consider industries with more than one firm, which leaves us with a cross-section of 180 SIC3 industries.

As stated above in Corollary 3, our model predicts compensation growth to be smaller in industries in which turnover is higher. As can be seen in Column 1 of the table, the cross-industries correlation between compensation growth and turnover is significantly negative. Again, we are not claiming causality, we are just observing that the correlation between endogenous variables observed in the data has the same sign as that predicted by the model. To give a sense of magnitudes, a one standard deviation increase in the level of turnover is associated with compensation growth that is lower by 8 percentage points. The result is robust to adding additional controls (see Column 2), namely the average market value of firms in the industry in 2010 ($AvgMktValue$) and the average market value growth between 2010 and 2016 ($AvgMktValueGrowth$).

As stated in Proposition 1, in equilibrium compensation increases simultaneously with complexity. Thus, empirically, across industries, large compensation growth should be associated with larger increase in complexity. This is the case in our data, both when complexity is proxied by the number of functions of an agent (Colums 3 and 4), and when it is proxied by the number of occurrences of the word complex (see Columns 5 to 8). To illustrate the economic magnitude of the effect, a one standard deviation higher increase in complexity (in the cross-section of industries) is on average associated with a compensation that is higher by 12 percentage points.

[18] As mentioned above, we think of this regression as an illustration of empirical patterns rather than a formal econometric test. In particular, we don't interpret the observed correlations in terms of causality.

7 Conclusion

We study how overlapping generations of agents acquire skills corresponding to more or less complex tasks. Complexity increases gross returns, as well as the cost of effort, which, in turn, increases agency rents. Because of the link between incentives and horizons, young and old generations are not perfect substitutes. Thus, young agents can choose more complex technologies, and correspondingly larger rents, than their elder peers. Competition, however, precludes large deviations from the choices of older generations. This leads to a progressive increase in complexity and rents. This key insight is robust to several changes in the modeling of the game between principals and agents, such as (1) giving more power to the principal (e.g. commit to simple technologies, access monitoring technologies, better search technologies) and (2) giving more bargaining power to the agent.

References

Almazan A., D. Chapman, K. C. Brown, and M. Carlson, 2004, "Why Constrain Your Mutual Fund Manager?" *Journal of Financial Economics*, 73: 289-321.

Axelson, U. and Bond, 2015, “Wall Street occupations: An equilibrium theory of overpaid jobs”, forthcoming *Journal of Finance*.

Baumol, W.J., 1990, “Entrepreneurship: Productive, unproductive and destructive.” *The Journal of Political Economy*, 5:893–921.

Biais, B., T. Mariotti, G. Plantin and J-C. Rochet, 2007, “"Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications", *Review of Economic Studies,* vol. 74: 345-390.

Becker, G. and G. Stigler, 1974, “Law enforcement, malfeasance, and compensation of enforcers,” *Journal of Legal Studies,* 3:1-18.

Bodenheimer, T., 2005, “High and rising health care costs; part 2: Technologic Innovation.” *Annals of Internal Medicine*, 932-937.

Bolton, P., and D. Scharfstein, 1990, “A theory of predation based on agency problems in financial contracting,” *American Economic Review*, 80, 94–106.

Bolton, P., T. Santos and J. Scheinkman, 2013, “Cream Skimming in Financial Markets,” Working Paper, Columbia University.

Brünnermeier, M., and M. Oehmke, 2009, Complexity in Financial Markets, Working paper, Princeton University.

Célerier, C. and B. Vallée, 2014, “What drives financial complexity? A look into the retail market for structured products”, Working paper, Harvard.

Carlin, B., 2009, “Strategic Price Complexity in Retail Financial Markets,” *Journal of Financial Economics*, 91, 278–287.

Chernenko, S., Hanson, S., Sunderam, A., 2014, “The Rise and Fall of Securitization”, Working paper, , Harvard Business School. .

DeMarzo, P.M., and M.J. Fishman, 2007, “Optimal Long-Term Financial Contracting,” *Review of Financial Studies*, 20, 2079–2128.

Diamond, P., 1971, “A Model of Price Adjustment,” *Journal of Economic Theory* 3, 156-168

Diamond, D., 1984, "Financial intermediation and delegated monitoring," *Review of Economic Studies*, 51, 393–414.

Edmans, A., X. Gabaix, T. Sadzik and Y. Sannikov, 2012, Dynamic CEO compensation, *The Journal of Finance, 67,* 1603-1647.

Glode, V., R. Green and R. Lowery, 2012, "Financial Expertise as an Arms Race," *The Journal of Finance*, 67, 1723-1759.

Greenwwod, R. and D. Scharfstein, 2012, "The Growth of Modern Finance," Working paper, Harvard Business School.

Grossman, S., and O. Hart, 1983, "An analysis of the principal–agent problem," *Econometrica*, 51, 7–45.

Henderson, B. and N. Pearson, 2011, "The dark side of financial innovation: A case study of the pricing of a retail financial product." Journal of Financial Economics, 100, 227–247.

Hochberg, Y., A. Ljungqvist, and Y. Lu, 2010, "Networking as a barrier to entry and the competitive supply of venture capital", *Journal of Finance*, 65, 829-859.

Holmstrom, B., 1979, "Moral Hazard and Observability", *The Bell Journal of Economics*, 10, 74–91.

Holmstrom, B., and J. Tirole, 1997, "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112, 663–691.

Kehoe, P. and D. Levine, 1993, "Debt–Constrained Asset Markets," *Review of Economic Studies*, 60, 865–888.

Kremer, M., 1993, "The O-Ring Theory of Economic Development", *The Quarterly Journal of Economics*, 108, 551-575.

Murphy, K., A. Shleifer, and R. Vishny,1991, "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics*, 106, 503–530.

Myerson, R., 2012, "A model of moral–hazard credit cycles," Working paper, University of Chicago.

New York Times, 2009, http://www.nytimes.com/2009/03/17/business/17sorkin.html?_r=1)

Oyer, P., 2008, "The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income," *Journal of Finance*, 63, 2601–2628.

Philippon, T. and A. Resheff, 2008, "Wages and Human Capital in the U.S. Financial Industry: 1909-2006", Working paper, NYU.

Rajan, R., 2008, "Banker's pay is deeply flawed", *Financial Times*, January 9.

Rajan, R. and L. Zingales, 1998, "Power in a Theory of the Firm", *Quarterly Journal of Economics*, 387-432.

Rampini, A. and S. Vishwanathan, 2010, "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance,* , 65, 2293–2322.

Rogerson, W., 1985, "Repeated Moral Hazard," *Econometrica*, 58, 69-76.

Philippon, T. and A. Resheff, 2008, "Wages and Human Capital in the U.S. Financial Industry: 1909-2006," Working Paper, New York University.

Sato, Y., 2015, "Opacity in Financial Markets," forthcoming *Review of Financial Studies*.

Townsend, R., 1979, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 22, 265–293.

Proofs:

Proof of Lemma 1: The Lagrangian is

$$\mathcal{L} = R(1+\rho) - w_t^t - \rho w_{t+1}^t + \mu_t(w_t^t + \beta w_{t+1}^t - b_t R) + \mu_{t+1}(w_{t+1}^t - b_t R),$$

where μ_t and μ_{t+1} are the multipliers of the time t and $t{+}1$ incentive constraints, respectively. The first order condition with respect to w_t is: $-1+\mu_t = 0$. Hence the incentive compatibility constraint at time t binds, i.e., $w_t^t + \beta w_{t+1}^t = b_t R$. The first order condition with respect to w_{t+1} is: $-\rho + \mu_t \beta + \mu_{t+1} = 0$. Substituting $\mu_t = 1$, $\mu_{t+1} = \rho - \beta$. When $\rho > \beta$, $\mu_{t+1} > 0$, so that the the incentive compatibility constraint at time $t+1$ also binds. Hence, the optimal compensation is as stated in the lemma.

QED

Proof that, under (10), shirking once is dominated by effort at both periods: Effort at both periods dominates shirking at t, followed by effort at $t+1$, if

$$R(b_t)[(1+\rho) - (1+\rho-\beta)b_t] \geq R(b_t)[(1-\Delta) + \rho(1-b_t)].$$

That is

$$\frac{\Delta}{1-\beta} \geq b_t. \tag{33}$$

Effort at both periods dominates effort at t, followed by shirking at $t+1$, if

$$R(b_t)[(1+\rho) - (1+\rho-\beta)b_t] \geq R(b_t)[(1-b_t) + \rho(1-\Delta)].$$

That is

$$\frac{\rho\Delta}{\rho-\beta} \geq b_t. \tag{34}$$

Now

$$\frac{1+\rho}{1+\rho-\beta} < \frac{1}{1-\beta} \leq \frac{\rho}{\rho-\beta},$$

Hence if (9) holds (which it does under (10)), then (33) and (34) also hold, so that effort at both periods is optimal.

QED

R and Z in the simple example: If $R(b) = \min[\alpha b + 1, \bar{R}]$,

$$Z(b) = (\alpha b + 1)[(1+\rho) - (1+\rho-\beta)b], \forall b \leq \frac{\bar{R}-1}{\alpha} \text{ and } Z(b) = \bar{R}[(1+\rho) - (1+\rho-\beta)b], \forall b > \frac{\bar{R}-1}{\alpha}.$$

Thus

$$Z^{'}(b) = \alpha[(1+\rho) - 2(1+\rho-\beta)b] - (1+\rho-\beta), \forall b \leq \frac{\bar{R}-1}{\alpha} \text{ and } Z^{'}(b) = -(1+\rho-\beta)\bar{R}, \forall b > \frac{\bar{R}-1}{\alpha}.$$

Thus, $Z' \leq 0$ for $b \geq \frac{\bar{R}-1}{\alpha}$. For $b \leq \frac{\bar{R}-1}{\alpha}$, $Z' \geq 0$ if and only if

$$\frac{\alpha(1+\rho) - (1+\rho-\beta)}{2(1+\rho-\beta)} \geq \alpha b.$$

This holds for all $b \leq \frac{\bar{R}-1}{\alpha}$, if

$$\frac{\alpha(1+\rho) - (1+\rho-\beta)}{2(1+\rho-\beta)} \geq \bar{R} - 1.$$

That is

$$\frac{1+\rho}{(1+\rho-\beta)} \geq \frac{2\bar{R}-1}{\alpha},$$

i.e., (11) holds. Hence, (11) implies $b^* = \frac{\bar{R}-1}{\alpha}$.

Proof of Lemma 2: Ω is non-empty because $b^* \in \Omega$. Indeed

$$Z(b^*) \geq \max[R(b^*_{t-1})(1-b^*_{t-1}) + \rho \max[R(b^*)(1-b^*), R(b^*_{t+1})(1-b^*_{t+1})] - \epsilon(1+\rho), Z(b^*_t) - \epsilon],$$

since $R(b)(1-b)$ is decreasing in b for $b \geq b^*$. Ω is compact. This compact subset of the real line has a unique maximum, b_t, which defines the unique solution of the maximisation program of the agent.

If $b_t \neq b_{max}$, it must be that $b_t < b_{max}$. If (20) did not hold, this would imply that the left–hand side of (20) would be strictly above its right-hand side. This strict inequality would by continuity extend to a neighbourhood of b_t included in $[0, b_{max}]$, which would contradict the fact that b_t is the maximum of Ω. So, either $b_t = b_{max}$, or b_t solves (20).

QED

Proof of Lemma 4: By Lemma 3, if b is to decrease between $t-1$ and t, i.e., $b^*_t < b^*_{t-1}$, we must have $b^*_{t+1} < b^*_t$. Then, as long as $b \leq b_{\max}$, (22) is

$$R(b^*_t)(1+\rho)(1-b^*_t) + \beta b^*_t R(b^*_t) = R(b^*_{t-1})(1-b^*_{t-1}) + \rho R(b^*_{t+1})(1-b^*_{t+1}) - \epsilon(1+\rho). \quad (35)$$

Denote $g(b_t) = R(b_t)(1 - b_t)$. In terms of g, (35) writes as

$$g(b_t^*)(1+\rho) + \beta b_t^* R(b_t^*) = g(b_{t-1}^*) + \rho g(b_{t+1}^*) - \epsilon(1+\rho).$$

That is

$$g(b_t^*) = \frac{g(b_{t-1}^*) + \rho g(b_{t+1}^*)}{1+\rho} - \frac{\beta b_t^* R(b_t^*)}{1+\rho} - \epsilon(1+\rho). \tag{36}$$

Since

$$\frac{g(b_{t-1}^*) + \rho g(b_{t+1}^*)}{1+\rho} - \frac{\beta b_t^* R(b_t^*)}{1+\rho} - \epsilon(1+\rho) < \frac{g(b_{t-1}^*) + \rho g(b_{t+1}^*)}{1+\rho},$$

we have

$$g(b_t^*) < \frac{g(b_{t-1}^*) + \rho g(b_{t+1}^*)}{1+\rho}.$$

By Jensen inequality (as g is concave and decreasing), this implies

$$b_t^* > \frac{b_{t-1}^* + \rho b_{t+1}^*}{1+\rho},$$

that is

$$b_t^* - b_{t+1}^* > \frac{1}{\rho}(b_{t-1}^* - b_t^*).$$

Because $\frac{1}{\rho} > 1$, This implies that, as t goes to infinity, $b_t^* - b_{t+1}^*$ goes to plus infinity, which, since $b_t^* \leq b_{\max}$, implies b_{t+1}^* goes to minus infinity, a contradiction since $b \geq 0$.

QED

Proof of Corollary 1: (23) rewrites as

$$\frac{R(b_t^*)}{R(b_{t-1}^*)} = (\frac{1 - b_{t-1}^*}{1 - (1-\beta)b_t^*}) - \frac{\epsilon(1+\rho)}{[1 - (1-\beta)b_t^*]R(b_{t-1}^*)}. \tag{37}$$

Proposition 1 implies that the left–hand side of (37) is larger than one. Hence (37) implies

$$(\frac{1 - b_{t-1}^*}{1 - (1-\beta)b_t^*}) - \frac{\epsilon(1+\rho)}{[1 - (1-\beta)b_t^*]R(b_{t-1}^*)} \geq 1,$$

which, in turn yields (24).

QED

Proof of Corollary 2: In the simple example, for $b \geq b^*$, (23) simplifies to

$$b_t^* = \frac{b_{t-1}^*}{1-\beta} + \frac{\epsilon(1+\rho)}{(1-\beta)\bar{R}}.$$

Thus, as ϵ goes to 0, we get (28).

QED

Proof of Lemma 5: $W(b)$ decreases with b, $\forall b > b^*$. $\forall b \leq b^*$, $W(b)$ increases with b if $(1+\beta)\alpha \geq Mc$. This is implied by our assumption that $R(b^*) \geq Mc$, if

$$(1+\beta)\alpha \geq R(b^*) = \bar{R}.$$

Now, since $\beta = \rho$, condition (11) simplifies to

$$\alpha(1+\beta) \geq 2\bar{R} - 1.$$

Hence, $W(b)$ increases with b if

$$\bar{R} \geq 1,$$

which holds. Hence, $W(b)$ increases with b, and $b^{**} = b^*$.

QED

Proof of Proposition 3: At time $t-1$, we had

$$b_{t-1}^* = (1+\beta)(1 - \frac{1}{\bar{R}}),$$

since the equilibrium choices of agents and principals were pinned down by the constraint that $b \leq (1+\beta)(1-\frac{1}{\bar{R}})$, and the employability constraint involved only variables set at time $t-1$ or $t-2$. The time t change in technology, however, affects the time t employability constraint, which becomes

$$\frac{\bar{R}}{\omega}[1+\beta-b_t^*] = \max[\frac{\bar{R}}{\omega}(1-(1+\beta)(1-\frac{1}{\bar{R}})) + \beta\frac{\bar{R}}{\omega}(1-b_t^*), 1+\beta]. \tag{38}$$

The left–hand side is the present value of the principal's gains if she hires the b_t^* agent. The right–hand side is the maximum of what the principal could get if i) hiring an old agent (with $b = (1+\beta)(1-\frac{1}{\bar{R}})$) at time t, and then another old agent (with $b = b_t^*$) at $t+1$, or ii) self investing.

At time $t-1$, the constraint that the principal be as well off hiring the agent as self investing was binding. The change in technology opens up the possibility that at time t this constraint becomes slack. In that case, (38) yields

$$b_t^* = \frac{1+\beta}{1-\beta}(1-\frac{1}{\bar{R}}) > b_{t-1}^*,$$

i.e., complexity starts growing again. This does not violate the constraint that the principal be better off hiring the agent than self investing if

$$\frac{\bar{R}}{\omega}[1+\beta-\frac{1+\beta}{1-\beta}(1-\frac{1}{\bar{R}})] \geq 1+\beta, \tag{39}$$

where the left-hand side is the present value of the gains of the principal hiring the b_t^* agent, while the right-hand side is the present value of self investment. (39) simplifies to

$$\bar{R}[1-\frac{1}{1-\beta}(1-\frac{1}{\bar{R}})] \geq \omega.$$

This holds for ω small enough iff

$$\beta\bar{R} \leq 1.$$

QED

Proof of Proposition 4: Once an agent has been selected (from the pair the principal was initially matched with), all is as in the benchmark model. The employability condition is still

$$b \leq b_{t-1}^* + \beta Min[b_t^*, b_{t+1}^*] + (1+\beta)\frac{\epsilon}{R}. \tag{40}$$

Prior to that stage, however, the agent also takes into account that her choice of b affects the probability to be chosen against his competitor. Rationally expecting the actions of the principals and the other agents, the agent expects to be hired with probability

$$\frac{2}{M}\Pr(b \leq b_t^* + \eta_t^{ji}) = \frac{2}{M}(1-F(b-b_t^*)).$$

So his expected rent is :

$$\begin{cases} \frac{2}{M}(1-F(b-b_t^*))bR & if\ b \leq b_{t-1}^* + \beta Min[b_t^*, b_{t+1}^*] + (1+\beta)\frac{\epsilon}{R} \\ 0 & \text{otherwise} \end{cases}$$

Denote by $b_c(b_t^*)$ the value of b at which $(1-F(b-b_t^*))b$ reaches its maximum. $b_c(b_t^*)$ is the solution of the implicit equation $b = \frac{1-F(b-b_t^*)}{f(b-b_t^*)}$ (uniquely defined by the monotone hazard rate assumption).

Without the employability constraint, if the agent anticipated his competitors to opt for b_t^*, he would choose $b_c(b_t^*)$. As long as $b_t^* < b_c(b_t^*)$, the employability constraint binds, and the agent can't choose $b_c(b_t^*)$. Rather he goes for b_t^*, just as in the benchmark case of Proposition 1.

Now suppose that i) at time t, $b_t^* < b_c(b_t^*)$ holds, but ii) the subsequent complexity arising in the benchmark case, b_{t+1}^*, would be such that $b_{t+1}^* > b_c(b_{t+1}^*)$. Then, unlike in the benchmark case, equilibrium complexity goes to the fixed point of b_c

$$\frac{1-F(0)}{f(0)} = \frac{1}{2f(0)}.$$

Combining the two cases, the unique symmetric equilibrium complexity, b_t^*, is defined recursively by:

$$\begin{cases} b_t^* = \frac{b_{t-1}^*}{1-\beta} + \frac{1+\beta}{1-\beta}\frac{\epsilon}{R} & if\ \frac{b_{t-1}^*}{1-\beta} + \frac{1+\beta}{1-\beta}\frac{\epsilon}{R} < min[b_{max}, \max[b^*, \frac{1}{2f(0)}]] \\ min[b_{max}, \max[b^*, \frac{1}{2f(0)}]] & \text{otherwise} \end{cases}$$

QED

Definition of the variables used in the empirical analysis

Variables used in individual regressions:

- $AvgTotalCompensation_{i,f}$: Average Total Annual Cash Compensation between 2010 and 2016 of individual i while working at firm f.
- $NbFunction_{i,f}$: Number of function of individual i at firm f in 2010.
- $NbComplex10K_f$: Number of occurrence of the string "complex" in the 10-K filing in 2010 of the company f employing the individual.

Variables used in industry regressions:

- $AvgTotalCompensationGrowth_j$: Growth of the average compensation in industry j, i.e., (Average Total Annual Cash Compensation in 2016)/(Average Total Annual Cash Compensation in 2010), using all individuals working at a firm in industry j in 2010 and/or 2016.
- $\Delta NbFunction_j$: Variation in the average number of function of individuals in industry j between 2010 and 2016, i.e., (Average number of function in 2016)-(Average number of function in 2010).
- $\Delta NbComplex10K_j$: Variation, between 2010 and 2016, in the average number of occurrence of the string "complex" in the 10-K filings of companies belonging to industry j, i.e., (Average number of occurrence in 2016)-(Average number of occurrence in 2010).
- $Turnover_j$: Average yearly turnover of firms in industry j between 2010 and 2016, where the yearly turnover is defined as the ratio between the number of workers leaving the firm in the year and the total number of workers this year.
- $AvgMktValue_j$: Average market value of firms in industry j in 2010.
- $AvgMktValueGrowth_j$: Growth of the average market value of firms in industry j between 2010 and 2016, i.e., (Average Market Value in 2016)/(Average Market Value in 2010) - 1.

	log($AvgTotalCompensation$)		
	(1)	(2)	(3)
Intercept	13.56***	13.84***	13.49***
	(0.02)	(0.32)	(0.02)
$NbFunction$	0.01*	0.02***	-
	(0.005)	(0.004)	
$NbComplex10K$	-	-	0.01**
			(0.004)
Firm F.E.	-	Yes	-
R^2	0.0003	0.43	0.002
Nb. Obs.	9,724	9,724	15,425

Table 1: Regression of the log of the average compensation ($AvgTotalCompensation$) on the number of functions ($NbFunction$) and the number of occurrence of "complex" in the employer's 10-K filings ($NbComplex10K$), between 2010 and 2016, at the individual level. ***, ** and * correspond to rejection of the null that the coefficient is zero, respectively at the 1%, 5% and 10% level. Standard errors are reported in parentheses and are clustered at the firm level for the regression in Column (3). Observations lying above (below) the quantile 99% (1%) of the dependant or independent variable distribution are removed.

$AvgTotalCompensationGrowth$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	1.41***	1.43***	1.23***	1.25***	1.18***	1.20***	1.37***	1.38***
	(0.10)	(0.10)	(0.03)	(0.03)	(0.03)	(0.03)	(0.10)	(0.10)
$Turnover$	-1.89**	-1.91**	-	-	-	-	-1.57*	-1.59*
	(0.85)	(0.82)					(0.89)	(0.89)
$\Delta NbFunction$	-	-	0.27***	0.26***	-	-	0.24***	0.23***
			(0.07)	(0.07)			(0.06)	(0.07)
$\Delta NbComplex10K$	-	-	-	-	0.03**	0.03**	0.02*	0.02**
					(0.01)	(0.01)	(0.01)	(0.01)
Controls	-	Yes	-	Yes	-	Yes	-	Yes
R^2	0.02	0.05	0.07	0.09	0.03	0.05	0.11	0.13
Nb. Obs.	180	180	180	180	180	180	180	180

Table 2: Regression of the growth in the average compensation ($AvgTotalCompensationGrowth$) on average yearly turnover ($Turnover$), variation in the average number of functions ($\Delta NbFunction$) and change in the number of occurrence of the string "complex" in companies' 10-K filings ($\Delta NbComplex10K$), between 2010 and 2016, at the industry level. Controls include the average market value of firms in the industry in 2010 ($AvgMktValue$) and the average market value growth between 2010 and 2016 ($AvgMktValueGrowth$). Variable definitions are provided above. ***, ** and * correspond to rejection of the null that the coefficient is zero, respectively at the 1%, 5% and 10% level. Standard errors are reported in parentheses. Observations corresponding to industries containing only one firm or lying above (below) the quantile 99% (1%) of the dependant variable distribution are removed.