Executive Compensation: A General Equilibrium Perspective* †

Jean-Pierre Danthine
Member of the Governing Board of the Swiss National Bank

John B. Donaldson
Graduate School of Business, Columbia University

July 6, 2010

Abstract

We study the dynamic general equilibrium of an economy where risk averse shareholders delegate the management of the firm to risk averse managers. The optimal contract has two main components: an incentive component corresponding to a non-tradable equity position and a variable "salary" component indexed to the aggregate wage bill and to aggregate dividends. Tying a manager's compensation to the performance of her own firm ensures that her interests are aligned with the goals of firm owners and that maximizing the discounted sum of future dividends will be her objective. Linking managers' compensation to overall economic performance is also required to make sure that managers use the appropriate stochastic discount factor to value those future dividends. General equilibrium considerations thus provide a potential resolution of the "pay for luck" puzzle. We also demonstrate that one sided "relative performance evaluation" follows equally naturally when managers and shareholders are differentially risk averse.

JEL classification: E32, E44
Keywords: incentives, optimal contracting, stochastic discount factor, pay-for-luck, relative performance

*Portions of this paper have circulated under the title: The Macroeconomics of Delegated Management. We thank Christian Hellwig, Erwan Morellec, and Deszoe Salay and seminar participants at Wharton, IMF, Universities of Washington, Santa Barbara, Mannheim, Cergy Pontoise and Ecares-Brussels for their comments. The research was largely conducted when Danthine was affiliated with the Université de Lausanne, the Swiss Finance Institute, and CEPR. Donaldson's work has benefited from financial support of the Faculty Research Fund, Graduate School of Business, Columbia University. Danthine's research was carried out within the National Center of Competence in Research "Financial Valuation and Risk Management." Corresponding Author: John B. Donaldson, School of Business, Columbia University, 3022 Broadway, New York, New York 10027, Tel. 212.854.4436, jd34 @columbia.edu

†The views expressed in this paper are those of the authors and do not necessarily represent those of the Swiss National Bank.

1. Introduction

We construct a dynamic general equilibrium model with agency conflicts between risk averse shareholders and managers and derive the first best compensation contract. General equilibrium considerations impose properties on this contract that, in turn, bring a natural resolution to two outstanding anomalies in the executive compensation literature. These "puzzles" are introduced below.

Standard incentive theory suggests that managerial incentive pay should depend on events exclusively under the manager's control, and not on events exogenous to her own efforts or decisions. To the extent that significant components of a CEO's pay are directly or indirectly (via, e.g., option grants) related to her own firm's stock returns, this theory advocates compensation mechanisms that reward the CEO only to the extent that her own firm's stock outperforms a chosen sectoral or broad based market index. Accordingly, such considerations mandate "relative performance evaluation" or RPE, that is, a negative relationship between a CEO's compensation and the chosen comparative performance benchmark. Surprisingly, RPE contracts are not commonly observed: large components of the average CEO's compensation appear to be positively related to factors totally beyond her control. This is the "pay-for-luck" puzzle.

Garvey and Milbourne (2006) further refine the puzzle by observing a stronger link of executive pay to market returns when returns are positive than when they are negative: executives enjoy good luck and appear to be insulated from bad luck, a practice referred to as "one sided RPE." This asymmetry in deviations from RPE constitutes the second puzzle.

Our explanation for these diverse anomalies will proceed along the following lines. Internal incentive considerations imply that the manager's compensation must be tied to the performance of her own firm. Performance-based compensation ensures that her interests are

appropriately aligned with the goals of the firm's owners and that maximizing the sum of future dividends will be her objective. Correct intertemporal decision-making, however, also requires that managers use the appropriate stochastic discount factor to value those future dividends. If shareholders are well diversified and exposed only to aggregate risks, aligning managers' and shareholders' stochastic discount factors (or SDFs) will typically require tying the managers' remuneration to aggregate state variables as well.

We make this point in a fully explicit context where the typical firm owner is the representative agent of the standard macroeconomic model. If managers and shareholders are equally risk averse, managers' decisions will be correct from the shareholders' perspective if and only if managers' and shareholders' consumption streams are in direct proportion to one another. This implies that managers must at least in part be compensated in proportion to the aggregate wage bill and to the payouts from all other firms. In other words, general equilibrium considerations, per se, demand a partial renunciation of RPE; as such, they argue directly for "pay for luck" as it is commonly understood and observed. We further demonstrate that one sided "pay for luck" is, in large measure, a consequence of optimal contracting in general equilibrium when managers and shareholders are differentially risk averse.

Ours is a simple infinite horizon dynamic general equilibrium model where both shareholders and managers are risk averse. The advantage of our set-up is that we can identify the contract that implements the first best allocation and, as a consequence, be fully specific as to its requirements. Reality is likely to be more murky, in particular because firm ownership may deviate from our representative agent assumption and because firm owners' information on managers' private wealth and actions may be incomplete. Yet, the essential lessons drawn from our simple set-up remain broadly applicable. To the best of our knowledge our model is a first

application of dynamic agency theory in a world where both principal and agent are risk averse.

An outline of the paper is as follows. Section 2 spells out the model. Section 3 characterizes the first best allocation of resources. Section 4 argues that there exists a contract decentralizing the first best allocation of resources, and completely characterizes this contract under the assumption that the manager's effort level is immaterial for production. The first best contract requires not only endowing the manager with a non-tradable equity share of the firm but also ensuring that the time series properties of the manager's stochastic discount factor, and thus her consumption, are identical to those of the firm owners. This latter condition in turn requires that the manager's remuneration includes a time-varying salary-like component whose properties are indexed to the aggregate wage bill. Section 5 generalizes this characterization to the situation where the manager's (unobservable) effort is essential for production while Section 6 develops the case of an economy with multiple firms. The salary component of a manager's remuneration must then include a share in the aggregate economy-wide dividend payment but, rather strikingly, a manager's equity stake in the firm under management cannot exceed her share of the market portfolio. Section 7 explores whether bond trading might be an adequate substitute for the first-best contract while Section 8 details the related literature. Section 9 concludes.

2. The model economy

For ease of exposition we start with the assumption that the entire economy's output is produced by a single perfectly competitive firm. Section 6 discusses the extension to many firms. There is a continuum of identical agents of measure $(1+\mu)$, a subset of which -- of measure μ -- is selected at the beginning of time to manage the firm permanently. The rest act as workers and shareholders. Managers are self-interested and are assumed to make all the relevant decisions in view of maximizing their own intertemporal utility. When they make the hiring and investment

decisions on behalf of firm owners, managers are viewed as acting collegially and thus we may refer to them collectively as "the manager."

At the center of our attention is the repeated principal-agent problem between the (risk averse) shareholders of the firm and the (risk averse) manager and its general equilibrium dimension. This agency problem has two distinct features. One is the familiar moral hazard dimension: the executive's effort choice is non-verifiable. The other, more important aspect is one of asymmetric information: the manager possesses specific knowledge of the firm's operation that is not available to shareholders. One of the main motives for delegation is, indeed, to relieve shareholders of the day-to-day operation of the firm and the information requirements it entails. This means that shareholders delegate to the manager the hiring and investment decisions and all that goes with them (human resource management, project evaluation, etc.) but that, as a by-product, they lose the informational base upon which to evaluate and monitor the manager's performance.

The representative **shareholder-worker-consumer** is confronted with a work/leisure decision and a portfolio investment decision. The form of his optimization problem is standard. The representative shareholder-worker's problem reads:

$$V^s(\hat{s}_0) = \max_{\{c_t^S, n_t^S, z_{t+1}\}} E\sum_{t=0}^{\infty}\beta^t[u(c_t^s) - \hat{H}(n_t^s)] \tag{1}$$

$$s.t.$$

$$c_t^s + q_t z_{t+1} \leq \left(q_t + d_t\right) z_t + w_t n_t^s,$$

$$c_t^s, z_t, n_t^s \geq 0, \forall t;$$

$$\hat{s}_{t+1} \sim dG\left(\hat{s}_{t+1}; \hat{s}_t\right), \hat{s}_0 \text{given.}$$

In problem (1), $u(\cdot)$ is the consumer-worker-investor's (homogeneous) period

utility-of-consumption function, $\hat{H}(\cdot)$ is his disutility of work function, c_t^s his period t consumption, n_t^s his period t labor supply and z_{t+1} the fraction of the single equity share purchased by him at the end of period t, while $G(.)$ describes the transition probabilities for the relevant state variables. As for prices, w_t denotes the competitive equilibrium wage rate, and q_t the ex-dividend (d_t) share price.

The firm is fully described by a constant returns to scale production function $f(k_t, n_t, \mu e_t)\lambda_t$ where k_t is capital stock available at the beginning of period t, n_t stands for employment, e_t is per capita managerial effort, and λ_t is the customary aggregate technology shock. The law of motion for capital stock is $k_{t+1} = (1-\Omega)k_t + i_t$ where i_t is investment and Ω is the rate of depreciation.

At the beginning of period t, **the manager** privately observes the realization of the productivity parameter λ_t; she then makes her utility-maximizing decisions (c_t^m, e_t, n_t, i_t) in light of her remuneration contract, $g^m(x_t, \hat{s}_t)$. Here c_t^m is the manager's period t consumption while x_t is a measure of the firm's performance to be identified later. Managerial remuneration may also depend on other economic variables observable by the firm owners (and on which they may write contracts). We denote by $\hat{s}_t$ the state of the economy as perceived by firm owners while s_t represents the true state of the economy as perceived by the fully-informed manager: $\hat{s}_t$, in particular, differs from s_t in that it does not include λ_t since the latter is the private information of the manager. We make precise the manager's informational advantage shortly. The manager is not given access to capital markets and she has no outside source of income or private wealth.[1]

[1] In Section 7 we explore the consequences of letting the manager trade a risk free bond.

She therefore consumes the income she receives from the firm. This assumption is essential to identifying, unambiguously, the first best contract. Given that the contract we discuss leads to the first best allocation, the assumption, in effect, is not restrictive. Our analysis can be extended without difficulty to situations where firm owners have full information on the manager's outside income and actions.

Given a level of effort e_t, decisions (n_t, i_t) yield distributions or dividends

$$d_t = f(k_t, n_t, \mu e_t)\lambda_t - n_t w_t - \mu g^m\left(x_t, \hat{s}_t\right) - i_t \equiv \hat{d}_t - \mu g^m\left(x_t, \hat{s}_t\right),$$

where $\hat{d}_t$ is free-cash-flow before payment to managers.

Let $u(\cdot)$ be a homogeneous function representing the manager's preferences over consumption, $H(\cdot)$ her disutility of effort, β the discount factor common to all economic agents and $F(\cdot)$ the probability transition function on λ_t. The manager's problem then reads:

$$V^m\left(s_0\right) = \max_{\left\{n_t, i_t, c_t^m, e_t\right\}} E\sum_{t=0}^{\infty}\beta^t\left[u\left(c_t^m\right) - H(e_t)\right] \tag{2}$$

s.t.

$$c_t^m = g^m\left(x_t, \hat{s}_t\right),$$

$$x_t = x(i_t, n_t, e_t; k_t, \lambda_t),$$

$$k_{t+1} = \left(1-\Omega\right)k_t + i_t, k_0 \text{ given}$$

$$V^m(s_0) \geq V^s(\hat{s}_0) \tag{3}$$

$$c_t^m, e_t, i_t, n_t \geq 0,$$

$$s_{t+1} \sim dF\left(s_{t+1}; s_t\right); s_0 \text{ given.}$$

Note that we assume both agent types have the same discount factor and the same

preferences over consumption[2]. The potential conflict of interests between the two agent classes - to be described shortly - arises endogenously and is not a result of postulated differences in preferences (in contrast with much of the literature - see Section 7). Constraint (3) guarantees the willingness of those shareholder-workers selected to manage the firm to assume that role (managers' participation constraint). Under the optimal contract, it will automatically be satisfied in equilibrium and thus we suppress it going forward (see footnotes 12 and 15).

We conclude this section by making explicit the information sets of, respectively, the shareholder-worker and the manager:

$$I_t^s = \{(q_\tau, w_\tau, d_\tau, z_\tau) : \tau \le t\}, \text{with } \hat{s}_t = (q_t, w_t, d_t, z_t),$$

and

$$I_t^m = \{(q_\tau, w_\tau, d_\tau, k_\tau, \lambda_\tau, \overline{z}_\tau \equiv 1) : \tau \le t; \Omega, f(\cdot, \cdot, \cdot)\}, \text{with } s_t = (q_t, w_t, d_t, k_t, \lambda_t).$$

Here $\overline{z}_t$ denotes the total number of shares outstanding. Since it is a time invariant quantity we elect to suppress it notationally. Moreover the manager is obviously aware of the form of the production function $f(\cdot, \cdot, \cdot)$ and the rate of capital depreciation Ω. Knowledge of these quantities by the shareholders is not assumed and, in fact is unnecessary for them to have rational expectations, $dG(\cdot;\cdot)$, over the stationary dividend, wage and asset price series (they operate, essentially, in a Lucas (1978) economy). With regard to state variables, the essential information advantage enjoyed by the manager is knowledge of the productivity shock λ_t. With this in mind, we sometimes write $V^m(k_t, \lambda_t)$ as a notational stand-in for $V^m(s_t)$ in order to emphasize the distinction. Lacking information as to the precise form of the production technology, the (possibly uncertain) depreciation rate, and aggregate employment, shareholder-workers are not in a position

[2]We consider the implication of differences in risk aversion between managers and shareholders at the end of Section 4.

to recover s_t from $\hat{s}_t$ and thus to monitor directly the investment and employment decisions of the manager.[3] This follows from the fact that neither of the manager's decisions, i_t and n_t, is observable, but only their consequence after the realization of the shock, in the form of d_t and w_t. But, without knowing k_t, the form of $f(\cdot,\cdot,\cdot)$ or the aggregate employment level, λ_t cannot be recovered from the pair (d_t, w_t).[4]

At this stage we view the manager's contract as having been imposed on the economy by a social planner but will subsequently argue (Section 4) that it has the same form as would evolve in a standard contracting setting where the $g^m(x_t, \hat{s}_t)$ and $x_t = x(\hat{s}_t)$ functions are explicitly chosen by the shareholder-workers. Accordingly, equilibrium is defined in the straightforward way:

<u>Equilibrium</u>: For a given managerial contract $g^m(x_t, \hat{s}_t)$ and performance measure $x_t = x(\hat{s}_t)$, equilibrium in the economy defined by problems (1) and (2) – (3) is a pair of price functions $w_t = w(s_t)$ and $q_t = q(s_t)$ such that:

(i) $c_t^s = c(\hat{s}_t)$, $n_t^s = n^s(\hat{s}_t)$ and $z_t = z(\hat{s}_t)$ solve problem (1);

(ii) $n_t = n(s_t)$, $c_t^m = c(s_t)$, $i_t = i(s_t)$ and $e_t = e(s_t)$ solve problem (2) – (3); and

(iii) markets clear:

$$c_t^s + \mu c_t^m + i_t = f(k_t, n_t)\lambda_t$$

$$n_t^s = n_t$$

[3]In particular, a contract which remunerates the manager if she takes the decisions in the best interest of shareholders, and harshly penalizes her otherwise is not feasible.

[4]In a rational expectations equilibrium such as the one studied here, $q_t = k_{t+1}$ and "Tobin's Q" is always one. By observing the equity price, shareholders can thus identify the capital stock level. The addition of either costs of adjustment to capital or distinguishing between tangible or intangible capital in the production technology breaks this equivalence. Both generalizations are accommodated in our setting. Nevertheless, in the present setting with knowledge of the capital stock but without knowledge of $f(\cdot,\cdot,\cdot)$ or n_t, λ_t cannot be inferred from the wage, equity price, and dividend series.

$$z_t = 1$$

While this sense of an equilibrium accommodates any contract, we are interested in the managerial contract which guarantees a 1st-best allocation.

3. Characterizing the first best allocation

In this section we characterize the first best allocation for the economy of Section 2. Noting that the aggregate state of the economy is given by (k_t, λ_t), the central planner's problem is as follows:

$$\max_{\{n_t, i_t, c_t^m, c_t^s, e_t\}} E\sum_{t=0}^{\infty}\beta^t \left\{\mu M[u\left(c_t^m\right) - H(e_t)] + (1-M)[u(c_t^s) - \hat{H}(n_t^s)]\right\} \tag{4}$$

s.t.

$$\mu c_t^m + c_t^s + i_t = f(k_t, n_t, \mu e_t)\lambda_t,$$

$$k_{t+1} = \left(1-\Omega\right)k_t + i_t, k_0 \text{ given,}$$

$$c_t^m, c_t^s, e_t, i_t, n_t \geq 0,$$

$$\lambda_{t+1} \sim dF\left(\lambda_{t+1}; \lambda_t\right), \lambda_0 \text{ given,}$$

where M and $1-M$ are arbitrary welfare weights assigned to an individual manager and shareholder, respectively. We introduce the following assumptions:

A.1: $u(\cdot)$ is twice continuously differentiable, strictly concave and increasing on R^+ and homothetic; the Inada conditions hold.

A.2: $\hat{H}(\cdot)$ is twice continuously differentiable, strictly convex and increasing on R^+.

A.3: $f(\cdot,\cdot)$ is twice continuously differentiable, strictly concave and increasing on $R^+ \times R^+$; the Inada conditions hold and $dF(\ ;\ )$ satisfies the Feller Property.

A.4: $H(\cdot)$ is twice continuously differentiable, strictly increasing and convex on R^+.

A standard result follows.

Theorem 1. Suppose A.1-A.4 hold. Then there exist a differentiable value function $W(k_t, \lambda_t)$ and continuous policy functions $\{n(k_t, \lambda_t), e(k_t, \lambda_t), i(k_t, \lambda_t), c^m(k_t, \lambda_t), c^s(k_t, \lambda_t)\}$ which solve problem (4). Furthermore, there exist $\{\underline{k}, \overline{k}\}$ such that $\underline{k} \le k_t \le \overline{k}, \forall t$ provided $k_0 \in [\underline{k}, \overline{k}]$.

The recursive representation of problem (4) is

$$W(k_t, \lambda_t) = \max_{(e_t, i_t, n_t^s, c_t^m)} \{\mu M[u\left(c_t^m\right) - H(e_t)]$$

$$+(1-M)[u\left(f(k_t, n_t^s, \mu e_t)\lambda_t - i_t - \mu c_t^m\right) - \hat{H}(n_t^s)]$$

$$+\beta \int W\left((1-\Omega)k_t + i_t, \lambda_{t+1}\right) dF(\lambda_{t+1}; \lambda_t)\}. \tag{5}$$

Under A.1-A.4, the necessary and sufficient F.O.Cs for (5) are, $\forall t$,

$$u_1(c_t^s) f_2(k_t, n_t^s, \mu e_t)\lambda_t = \hat{H}_1(n_t^s), \tag{6}$$

$$1 = \beta \int \frac{u_1(c_{t+1}^m)}{u_1(c_t^m)} \left[f_1(k_{t+1}, n_{t+1}, \mu e_{t+1})\lambda_{t+1} + (1-\Omega)\right] dF(\lambda_{t+1}; \lambda_t), \tag{7}$$

$$(1-M)u_1(c_t^s) = Mu_1(c_t^m), \text{implying} \frac{u_1(c_{t+1}^s)}{u_1(c_t^s)} = \frac{u_1(c_{t+1}^m)}{u_1(c_t^m)}, \tag{8}$$

$$(1-M)u_1(c_t^s) f_3(k_t, n_t, \mu e_t)\lambda_t = MH_1(e_t), \tag{9}$$

and

$$\mu c_t^m + c_t^s + i_t = f(k_t, n_t, \mu e_t)\lambda_t \equiv y_t. \tag{10}$$

Using (8), condition (9) can be written

$$u_1(c_t^m) f_3(k_t, n_t, \mu e_t)\lambda_t = H_1(e_t). \tag{11}$$

Condition (6) is the standard marginal condition determining the worker's optimal supply of labor. Condition (11) is the equivalent condition for the effort level of the manager. Equation (7)

is an equally standard condition determining investment. Note that the relevant intertemporal rate of substitution is the manager's, but the Pareto Optimal income and risk sharing condition (8) implies that this could equally well be the shareholder's. Finally equation (10) is the overall resource availability constraint. In the next section we discuss the optimal contract under the assumption that the effort of the manager plays no role in determining the period output of the firm. This specialized case allows us not only to present the intuition underlying our more general contracting context but also to relate our results directly to the standard macroeconomics literature.

4. A first-best contract: the no effort case

For ease of exposition we take as a benchmark the situation where the effort of the manager is irrelevant, yet where the manager is offered a performance-based contract, and concentrate on the more essential problem resulting from asymmetric information.[5] We assume $f_3 \equiv 0$ and drop the manager's effort from the production function for notational simplicity. In that context we demonstrate the following:

Theorem 2. Suppose A.1-A.4 are satisfied and the manager's effort is immaterial to production. Then there exists a first best contract, $g^m(x_t, \hat{s}_t) = \varphi \hat{d}_t + \varphi w_t n_t, \varphi > 0$, under which the competitive equilibrium delivers the first best allocation of resources. The contract is unique up to the arbitrary parameter φ.[6,7]

Proof. Theorem 2 contains the main idea of this paper. We demonstrate it step by step,

[5]We acknowledge that under this hypothesis a fixed remuneration would leave the manager's welfare unaffected by her investment and hiring decisions. It could then be assumed that she would be truthful and adopt the policies advocated by shareholders. As is intuitively clear and will be emphasized later, however, this indifference is broken and some link between compensation and performance made necessary as soon as managers' effort matters. We anticipate this more realistic context in the simple setting of this section.

[6]In all our model constructs, the first-order conditions are necessary and sufficient under maintained assumptions A.1 -A.4. Providing a contract under which the first-order conditions for the Pareto Optimum coincide with those of the competitive equilibrium is thus sufficient to guarantee these two economic constructs have identical properties. We appeal to this logic throughout the paper.

[7]The parameter M identifies φ uniquely.

with some discussion, so as to emphasize the roles of the various components of the first-best contract.

The representative shareholder's problem (1) has the following recursive representation

$$V^s(z_t,\hat{s}_t) = \max_{\{z_{t+1},n_t^s\}} \{u\left(z_t\left(d_t+q_t\right)+w_t n_t^s - q_t z_{t+1}\right) - \hat{H}\left(n_t^s\right) \tag{12}$$

$$+\beta\int V^s\left(z_{t+1},\hat{s}_{t+1}\right)dG\left(\hat{s}_{t+1};\hat{s}_t\right)\}$$

whose solution is characterized by:[8]

$$u_1(c_t^s)w_t = \hat{H}_1(n_t^s), \tag{13}$$

$$u_1(c_t^s)q_t = \beta\int u_1(c_{t+1}^s)[q_{t+1}+d_{t+1}]dG\left(\hat{s}_{t+1};\hat{s}_t\right). \tag{14}$$

Note, from (13), that worker-shareholders' (static) labor supply decisions are independent of the probability distribution summarizing their information. The same cannot be said of their portfolio investment decisions (equation (14)) which forms the basis for equity pricing. Observe that no information beyond what shareholders possess can be included in the stock price, so that the stock market is not strongly informationally efficient in the sense that the stock's price is not a sufficient statistic for the information held by insiders (the managers).

Under appropriate conditions, the manager's problem (2) has recursive representation:[9]

---

[8]It follows from Blackwell's (1965) Theorem and the results in Benveniste and Scheinkman (1979) that a differentiable, bounded $V^s(\cdot)$ satisfying (12) exists together with unique policy functions characterized by (13) and (14) provided $u(\cdot)$ and $\hat{H}(\cdot)$ are increasing, continuously differentiable and concave, $q(\cdot)$ and $w(\cdot)$ are continuous, and that $dG(\cdot)$ has the property that it is continuous and whenever $h(d,q,w)$ is continuous, $\int h(d',q',w')dG(d',q',w';d,q,w)$ is continuous as a function of (d,q,w). The continuity of $q(\cdot)$ and $w(\cdot)$ is then confirmed in equilibrium.

[9]It again follows from Blackwell's (1965) Theorem and the results in Benveniste and Scheinkman (1979) that a differentiable, bounded $V^m(\ )$ exists that solves (15) provided $u(\cdot)$ and $f(\cdot)$ are increasing, continuous and bounded, and that $g^m(\cdot)$ is itself continuous and that $dF(A',\lambda';A,\lambda)$ is also continuous with the property that for any continuous $h(k',A',\lambda')$, $\int h(k',A',\lambda')dF(A',\lambda';A,\lambda)$ is also continuous in k and λ. In order for (16) and (17) to characterize the unique solution, the differentiability of $u(\cdot)$, $g^m(\cdot)$ and $f(\cdot)$ is required and

$$V^m\left(k_t,\lambda_t\right)=\max_{\{i_t,n_t\}}\left\{u\left(c_t^m\right)+\beta\int V^m(k_{t+1},\lambda_{t+1})dF(\lambda_{t+1};\,\lambda_t)\right\}. \tag{15}$$

The necessary and sufficient first order conditions to problem (15) can be written

$$u_1\left(c_t^m\right)g_1^m\left(x_t,\hat{s}_t\right)\frac{\partial x_t}{\partial n_t}=0, \tag{16}$$

$$-\frac{\partial x_t}{\partial i_t}=\beta\int\frac{u_1\left(c_{t+1}^m\right)g_1^m\left(x_{t+1},\hat{s}_{t+1}\right)}{u_1\left(c_t^m\right)g_1^m\left(x_t,\hat{s}_t\right)}\left[f_1\left(k_{t+1},n_{t+1}\right)\lambda_{t+1}+\left(1-\Omega\right)\right]dF\left(\lambda_{t+1};\,\lambda_t\right), \tag{17}$$

where this latter representation is obtained using a standard application of the envelope theorem.

Let us first hypothesize that $g^m(\cdot)$ is linear; i.e., $g^m(x_t,\hat{s}_t)=A_t+\varphi x_t$ where $A_t=A(\hat{s}_t)$ is independent of variables under the manager's direct control.[10] In what follows, we confirm that it is the unique first best contract. Given the linearity of the contract, $g_1^m(\cdot)$ is constant and for (7) to obtain from (17) it is necessary and sufficient that

$$\frac{\partial x_t}{\partial i_t}=-1.$$

Similarly, a comparison of equation (6) with (13) and (16) makes clear that for the standard optimality condition for employment to obtain given that g_1^m is constant, the measure of firm performance x_t must satisfy

---

$u(g^m(\cdot))$ must be concave. The assumptions made in this and the preceding footnote are maintained throughout the paper.

[10] That is, under the usual assumption that the unique firm is representative of a large number of identical firms behaving competitively.

$$\frac{\partial x_t}{\partial n_t} = \left[f_2(k_t, n_t)\lambda_t - w_t \right]$$

Integrating the two derivative conditions on x_t with respect to n_t and i_t, respectively, yields (up to a constant of integration):

$$x_t = f(k_t, n_t)\lambda_t - w_t n_t - i_t + \text{constant} \equiv \hat{d}_t + \text{constant}.$$

In other words, if there is to be no first-order distortion in the decisions of the manager, the manager's remuneration must be a linear function of the appropriate measure of firm performance and the latter must be free-cash-flow *before payments to managers*. Note that this identification corresponds to the minimal information requirement we may want to impose on worker-shareholders.

The last piece of the puzzle is to make sure that the optimal risk sharing condition (8) obtains. To discuss this issue, note that in equilibrium, at all dates t,

$$n_t^s = n_t, \tag{18}$$

$$z_t \equiv 1, \text{and} \tag{19}$$

$$y_t \equiv f(k_t, n_t)\lambda_t = c_t^s + \mu c_t^m + i_t \tag{20}$$

Equation (20) implies that condition (8) reads:

$$u_1(c_{t+1}^m) = u_1(\varphi \hat{d}_{t+1} + A_{t+1})$$

$$= u_1(y_{t+1} - i_{t+1} - \mu c_{t+1}^m) = u_1(c_{t+1}^s). \tag{21}$$

The homogeneity property of $u(\cdot)$, and thus $u_1(\cdot)$, in turn implies that equality (21) will be satisfied if the consumptions of the two agents are directly proportional to one another and thus to aggregate consumption, $c_t = y_t - i_t$. Consider the following equalities

$$c_t^m = \varphi \hat{d}_t + A_t$$

$$= \varphi[y_t - w_t n_t - i_t] + A_t$$

$$= \varphi(y_t - i_t) - \varphi w_t n_t + A_t \tag{22}$$

$$c_t^s = y_t - i_t - \mu c_t^m$$

$$= y_t - i_t - \mu[A_t + \varphi(y_t - i_t) - \varphi w_t n_t]$$

$$= (1 - \mu\varphi)(y_t - i_t) + \mu\varphi w_t n_t - \mu A_t. \tag{23}$$

These relations indicate that proportionality obtains if

$$A_t = \varphi w_t n_t.$$

It follows that $c_t^m = \varphi(y_t - i_t)$ and $c_t^s = (1 - \mu\varphi)(y_t - i_t)$ and equation (8) obtains.

It remains to argue that the assumption of contract linearity is not restrictive. This is the case since the optimal policy functions $c^m(\cdot)$, $c^s(\cdot)$, $i(\cdot)$ and $n^s(\cdot)$ arising from the solution to the Pareto problem (4) are unique. Since $g^m(x_t, \hat{s}_t)$, as we have defined it, replicates the optimal managerial consumption it must itself be unique and the original linearity assumption is not restrictive. This completes the demonstration of Theorem 2.[11] ■

Theorem 2 makes five assertions:

1. the appropriate measure of the firm's performance is $\hat{d}_t$, distributions before payment to managers;

2. the incentive element of the managers' remuneration is a fraction of distribution $\varphi\hat{d}_t$;

3. the optimal contract includes a "salary component" in addition to the incentive

[11] We argue for uniqueness in the following way. By the concavity of the objective function in problem (4), the associated policy functions $\{n(k_t, \lambda_t), i(k_t, \lambda_t), c^m(k_t, \lambda_t), c^s(k_t, \lambda_t)\}$ are all unique. Furthermore, by the argument above, $\hat{d}_t = \hat{d}(k_t, \lambda_t)$ is the unique aggregate on which incentive pay can be based. Since $c_t^m = \varphi\hat{d}_t + A_t, A_t = A(k_t, \lambda_t)$ is also unique and must equal $\varphi w_t n_t$. The parameter $\varphi = \varphi(M, \mu)$ is then uniquely determined by the optimality condition (8).

element;

4. the salary component is linearly related to the aggregate wage bill; and

5. the power of the incentive element, φ, also defines the exposure of the salary component to the aggregate wage bill.

The general message from this first approach may be summarized as follows. Contracting in general equilibrium requires not only aligning the "micro incentives" of managers and firm owners but also aligning their stochastic discount factors. To insure that the trade-offs internal to the firm are properly appreciated by the manager, it is appropriate to entitle her to a (non tradeable) equity position, hence to a claim to a fraction of present and future cash flows to capital. This will naturally guarantee that the manager will want to maximize the discounted sum of future expected dividends. In a multi-period world of risk averse agents, however, this is not sufficient. Shareholders want to ensure that the same stochastic discount factor as their own is applied by managers when tallying up future dividends. This is guaranteed by condition (21). This message is fully general. Under the assumptions of our model (homothetic utility functions and representative shareholder-worker), if the stochastic discount factors are to be aligned, the total compensation package of managers must be such that their consumption is proportional to aggregate consumption, that is, to $y_t - i_t = d_t + w_t n_t$. For this to be the case under the no-trading, no outside income assumption, the salary part of their remuneration must be the same fraction of the aggregate wage bill as the incentive portion is of $\hat{d}_t$. The overarching principle is that in order to select the investment and hiring policies the shareholder-workers would like, the manager must receive an income stream with identical characteristics. Since shareholder-workers receive the bulk of their income in the form of wage payments, the manager must as well.[12]

[12]We note that an appropriate choice of M will guarantee that the manager's welfare exceeds that of a representative

In the next section, we reinstate the effort dimension for completeness and show that the exposure to dividends (φ) will be determined by the extent of the moral hazard problem. But the central message will remain: in the general equilibrium of a world with a representative shareholder-worker, the various components of the manager's remuneration have to be "adapted" to one another in order to form an overall package that is proportional to aggregate consumption.

One interesting characteristic of this first-best contract is that it does not require the manager to communicate with the principal after observing the realization of the productivity shock. The first-best contract remains valid even if one interprets the signal λ_t as specific knowledge in the sense of Fama and Jensen (1983) and Jensen and Meckling (2002). There are a number of reasons for such an interpretation to be desirable. As emphasized by Jensen and Meckling (2002), knowledge transfer may involve costly delays. In addition, a particular value of λ_t and its implications for future productivity are meant to summarize a set of soft (in the sense of Stein (2003)) and continuously evolving elements of information on which it would be impossible or costly to base a compensation contract. These remarks are in the spirit of the information distinction made earlier.

Theorem 2 has the following interesting corollary:

Corollary (Equivalence Theorem). Suppose the manager is of measure $\mu = 0$ so that $\hat{d}_t = d_t$. Then under the linear contract $g^m(\hat{d}_t) = g^m(d_t) = \varphi w_t n_t + \varphi d_t$ the delegated management economy exhibits the same time series properties as, and is thus observationally equivalent to, the

---

shareholder and guarantee that the manager's participation constraint is satisfied. By equations (8) and (22)-(23) each M translates into a corresponding φ. Alternatively, we could also postulate a "prestige utility increment," ξ, associated with the office of manager. For any *M*, there will always be a $\xi = \xi(M)$ such that the welfare of the manager exceeds that of a shareholder-worker. The introduction of this constant term, furthermore, has no effect on any of the allocations or contract forms derived in this essay.

representative agent (real) business cycle model.[13]

The contribution of our Corollary is to extend the realm of application of the standard business cycle model. The measure zero assumption is made for convenience only to facilitate comparison with the standard representative agent model. To maintain comparability in the presence of a positive measure of managers, it would be necessary to increase the productivity of factors to make up for managers' consumption in a way such that the consumption level of shareholder-workers, and consequently their labor supply decision, remains unchanged in equilibrium.

The following additional information will prove useful when we come to discuss "one-sided" RPE contracts.

Theorem 3. Suppose the manager is of measure $\mu = 0$ and that the period utility functions of the manager and shareholder-worker are, respectively, $u(c_t^m) = \frac{(c_t^m)^{1-\gamma_m}}{1-\gamma_m}$ and $u(c_t^s) = \frac{(c_t^s)^{1-\gamma_s}}{1-\gamma_s}$, where γ_m, γ_s satisfy either $0 \le \gamma_s \le 1$ and $0 \le \gamma_m \le 1$ or $1 < \gamma_m$ and $1 < \gamma_s$. In this case the first-best contract is of the form $g^m\left(x_t, \hat{s}_t\right) = \left(\varphi w_t n_t + \varphi \hat{d}_t\right)^\theta$ where $\theta = \frac{1-\gamma_s}{1-\gamma_m}$. If $\left|\gamma_m - 1\right| < \left|\gamma_s - 1\right|$, the contract is convex.

Proof: See Appendix A.

The preceding discussion carries over unchanged. Note that the measure $\mu = 0$ assumption is not required. With differing risk aversions, however, the optimal allocations would

[13] As such this paper offers an alternative decentralization scheme to those of Prescott and Mehra (1980) and Brock (1982). The $\mu = 0$ assumption is required only to guarantee that the shareholder-workers have the same consumption level as the representative agent in an otherwise identical economy. Shorish and Spear (2005) propose an agency theoretic extension of the Lucas (1978) exchange model. They focus on asset pricing however.

no longer be fixed shares of output net of investment.[14]

At the conclusion of Section 3, we acknowledged that our sense of equilibrium simply postulated the existence of a particular contract, and then demonstrated that it led to a Pareto optimal allocation without deriving it from an expanded choice problem confronting the shareholder-worker; that is, $x_t = x(\hat{s}_t)$ and $g^m(x_t, \hat{s}_t)$ were not choice variables of the shareholder-worker's decision problem. Here we argue that the optimal contract that would be derived in this latter, more traditional way, must coincide with the contract proposed in Theorem 2. The logic is as follows. As noted in the proof of Theorem 2, homotheticity of the agents' utility function requires that, at all output levels, their relative consumption remains in fixed proportion. By this same property, for any allocation φ, either agent will choose the same investment plan conditional on the labor supply schedule. To say it differently, for any φ, managers choose the same investment policy as would the shareholder-workers were the latter group to manage the firm itself. No other contract can improve upon what is already a Pareto optimum. A formal derivation is provided in Donaldson et al. (2010).

In the next two sections we confirm the essential intuition obtained in Theorems 2 and 3 and extend the main result in two directions. First we deal with the case where managerial effort is a required input in the production process. In this case we show the essentials of the prior contract are preserved but the share parameter φ is no longer indeterminate, leading to the necessary inclusion of an additional term in the remuneration package. We subsequently relax the assumption of a single firm and identify the first best contract in a world with multiple firms subject to idiosyncratic risks.

[14] The conclusions to Theorems 2 and 3 – and Theorems 4 and 5 of the subsequent sections – are also preserved under model generalizations that allow stochastic depreciation and/or uncertainty in the marginal efficiency of investment ($k_{t+1} = (1-\Omega)k_t + h(i_t, \psi_t)$ where ψ_t is stochastic and $h(\quad)$ captures the net addition to capital stock resulting from investment).

5. The optimal contract: unobservable effort

We focus on the full formulation of our problem where the manager's effort is an essential element in the production process. Our main result is expressed in

Theorem 4. If A.1-A.4 are satisfied, then there exists a unique first best contract $g^m(x_t, \hat{s}_t)$ sufficient for the competitive equilibrium to deliver the first best allocation of resources. The optimal contract possesses the following distinguishing features:

(i) $g^m(x_t, \hat{s}_t) = \varphi \hat{d}_t + A_t$ where A_t does not depend on variables under the manager's control;

(ii) $\mu\varphi = 1$;

(iii) $A_t = \varphi w_t n_t + \xi(y_t - i_t)$, for some constant $\xi < 0$.

Proof: See Appendix A.

In substance the main difference with the case of the previous section is as follows: in order to obtain the first best effort level, the share of free-cash-flow to be awarded to managers is not indeterminate but must be such as to elicit the right level of effort. Depending on the cost of effort and on the role of managers in production, it may well be - as is the case in our (quite extreme) formulation - that the entire free-cash-flow must be awarded to them ($\mu\varphi = 1$). If this is the case, the general equilibrium condition uncovered in Section 4, which states that the salary component must be adapted to the incentive component of managers' remuneration, implies that managers should be entitled to the entire $\hat{d}_t$ *and* to the entire wage bill.

As is, we are led to the conclusion that the collective of managers should receive the entire value added which is of course not possible. But there is a way out. It consists in the portion of manager's remuneration which is exogenous to her own decisions being corrected by a term that is

negative and proportional to ($y_t - i_t$); that is, with $\xi < 0$. This is the essence of condition (iii). In words, the first-best contract stipulates that the managers' remuneration should change one-to-one with variations in $\hat{d}_t$ (without limited liability) with a salary component engineered to guarantee that, if (and only if) she takes on the first best decisions, her total remuneration will be proportional to aggregate consumption.[15, 16]

Note that we have cast the potential moral hazard problem in the extreme. If either effort is partially observable or if there is a maximum possible level of effort, then it is conceivable that the share of free-cash-flows to be allocated to managers will be significantly less than one and a contract close to the one derived in the previous section, with a salary component proportional to the wage bill but without the negative corrective term, may be feasible.

It is important to point out that this contract, while unusual in a macroeconomic setting, is a straightforward generalization of a classic contract discussed in the incentives literature. If the model is specialized by the absence of production uncertainty (no variation in the λ_t), the optimal contract is simply one in which the manager pays a fixed fee to run the firm after which he receives all the firm's output. Such an arrangement is the exact counterpart of the classic taxicab contract (see, e.g.,Lazear (1995), chapter 1). Under uncertainty the circumstances are only slightly more complex: the fee varies across states in a manner designed to provoke perfect risk sharing between the agent (the DM) and the principal (the shareholder-workers).

Another noteworthy feature of the contract is that it induces the manager to report cash flow to capital, $\hat{d}_t$, accurately (truth-telling). In particular, there is no incentive for the manager

[15]Re-introducing a participation constraint for the manager would lead to pinning down a single value for ξ and thus the level of manager's remuneration, not only its dynamic properties.

[16]The absence of the usual limited liability constraint explains why we obtain a first best contract under both moral hazard and asymmetric information . In equilibrium managers' remuneration if of course guaranteed to be positive. This is not the case off equilibrium.

to steal a portion of $\hat{d}_t$, in any period. Suppose the manager were to report a $\hat{d}_t^R$, where $0 < \hat{d}_t^R < \hat{d}_t$, while privately consuming the difference $(\hat{d}_t - \hat{d}_t^R)/\mu$ (per capita). In this case the aggregate consumption of the manager (retaining the optimal contract) is $\mu\varphi\hat{d}_t^R + \mu A_t + (\hat{d}_t - \hat{d}_t^R)$ which equals $\hat{d}_t + w_t n_t - \xi(y_t - i_t)$, the same as it would be under truth-telling. The simple intuition is as follows: since the manager effectively receives all the output while paying a variable fee to the shareholders, then stealing amounts to the manager stealing from herself, a pointless endeavor.

In Figure 1, we plot a representative sequence of A_t, $\varphi\hat{d}_t$ and total managerial compensation. As should be clear by now, the entire package is designed to generate a smooth consumption series for the manager, a series with the same intertemporal characteristics as the consumption series of the shareholder-worker. This outcome, however, requires that the manager adopt the optimal hiring and investment policies. Here we have arbitrarily decided that the class of managers has exactly the same consumption level as the class of shareholder-workers. The incentive portion of management's remuneration is $\varphi\hat{d}_t$. It is significantly more variable than the consumption series. The difference between the two series is A_t which appears almost to move one for one in the opposite direction as the $\hat{d}_t$ series, as must be the case if the sum of the two series is to acquire the required smoothness. Note, however, that the two series are determined independently: in particular the A_t portion of the remuneration is part of the manager's contract irrespective of the performance of her own firm. It is therefore clear that there will be many instances where the manager' salary component will compensate for the loss in her performance-based remuneration resulting from poor firm performance. Yet, the fact that the salary part of the remuneration depends on the aggregate state of the economy also means that

were the manager to deviate from the optimal hiring and investment decisions, a policy that would lead on average to a deterioration of her own firm's operating results, her own remuneration would be affected and (on average) fall below the first-best performance-based remuneration depicted here.[17]

[Insert Figure 1 here]

In general, the main message resulting from the general equilibrium dimension of our inquiry is that there must be a balance between the performance based and the non-performance based elements of the manager's remuneration. Given that the consumption series of the manager should be suitably smooth - because it must replicate the dynamic properties of the consumption of the representative shareholder - and that the measures of the firm's performance are bound to be highly variable, it is necessarily the case that the salary component of the manager's remuneration will more often than not cushion the impact of variation in performance-based remuneration on the manager's overall compensation package. This, we argue, is the rationale behind the "pay-for-luck" characteristics of managerial contracts.

Note that the most implausible aspect of the contract illustrated in Figure 1, that the salary component is always negative, is partly an artifact of our definition of the variables. Assume, indeed, that $\hat{d}_t$ takes values in an interval $[\hat{d}_{min}, \hat{d}_{max}]$; then the performance based component of the first-best contract could equally well be defined as $\mu\varphi(\hat{d}_t - \hat{d}_{min})$ while the salary component would be $A_t + \hat{d}_{min}$ which, depending on circumstances, may always be positive.

In the case where the manager and the shareholder are differentially risk averse, a result similar to Theorem 4 holds as well. Since it still must be the case that $\mu\varphi = 1$, the principal

[17] A short-lived manager could deviate and gain in the short run. Here we focus on permanent managerial contracts. In general, short term bias induced by short term contracts would have to be corrected by an average compensation package that would be rising over time.

difference will reside in the "giveback" portion of the A_t term. It will no longer be fixed as a proportion of $(y_t - i_t)$ but rather must be adapted, period by period, to satisfy the economy-wide budget constraint.

In concluding this section it is worth stressing that the optimal contract must be understood as one where the incentive component depends on firm level performance as measured by free-cash-flow, while the "salary" component depends on the *aggregate* wage bill. In the next section we formalize this distinction in a more realistic economy with many firms each with a separate manager.

6. Many firms

We now extend our analysis to the case of a large number J of atomistic firms. The management of each firm is of measure $\mu \neq 0$ and the total measure of the managerial class is μJ. Firm j is characterized by technology $f(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j$ on the basis of which it distributes

$$d_t^j = f(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j - n_t^j w_t - \mu g^j(d_t^j, s_t) - i_t^j = \hat{d}_t^j - \mu g^j(d_t^j, s_t).$$

Optimality conditions are straightforward generalizations of those obtained in Section 3, that is, $\forall t$ and $j = 1,..,J$,

$$u_1(c_t^s) f_2(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j = \hat{H}_1(n_t^s), \tag{24}$$

$$1 = \beta \int \frac{u_1(c_{t+1}^j)}{u_1(c_t^j)} \left[f_1(k_{t+1}^j, n_{t+1}^j, \mu e_{t+1}^j)\lambda_{t+1}^j + (1-\Omega) \right] d\hat{F}(\cdot,\cdot), \tag{25}$$

$$(1-M)u_1(c_t^s) = M u_1(c_t^j), \tag{26}$$

$$u_1(c_t^j) f_3(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j = H_1(e_t). \tag{27}$$

where c_t^j denotes the consumption of the manager of firm j and $g^j(\ )$ her contract. Our main result is

Theorem 5. If A.1-A.4 are satisfied, then there exists a set of contracts $g^j(x_t, \hat{s}_t), \forall j$, sufficient for the competitive equilibrium to deliver the first best allocation of resources. The optimal contracts possess the following distinguishing features:

(i) $g^j(x_t, \hat{s}_t) = \varphi \hat{d}_t^j + A_t^j$, where A_t^j does not depend on variables under the manager's control;

(ii) $\mu\varphi = 1,$;

(iii) $A_t^j = \varphi w_t n_t + \varphi d_t^{\tilde{j}} + \xi(y_t - i_t), \xi \ll 0,$ for $d_t^{\tilde{j}} \equiv \sum_{i \neq j} d_t^i$.

(iv) $c_t^j = c_t^m$, i.e., as a result of the above, all managers have the same consumption stream as required by equation (26).

Proof: See Appendix A.

Theorem 5 confirms the message of the previous two sections: aligning the interests of principal and agent in general equilibrium requires going beyond the typical conditions identified in partial equilibrium. Making sure that the managers do perceive the firm-internal trade-offs in the same way as firm owners is only the first step. Aligning the discount factors of the two agent types is the second. Using the same logic as before, this requires not only giving the manager a share of the aggregate wage bill but also a share of the aggregate stock market.

Furthermore, the manager's compensation must be as sensitive to the aggregate wage bill *and* to the aggregate dividend payment made by other firms (or, by approximation, to the economy's total GDP net of aggregate investment) as it is to the measure of performance of the firm she manages. Equivalently the optimal contract stipulates that *a manager's (direct or indirect) exposure to the equity value of the firm she manages should not exceed her exposure to the world market portfolio*. This prescription is very intuitive in the context of our model economy.

The presence of an effort dimension further results in the condition that the collective of managers must be exposed, at the margin, to the full increase in dividends resulting from their effort. These requirements together imply that they should be granted the entire world GDP! Hence, the necessity of a (negative) corrective term, which must be designed to preserve the fractional proportionality of the managers' consumption to aggregate consumption, arises.

One of the important lessons of our exercise is that being careful to align the stochastic discount factor necessarily means not tying the manager's remuneration exclusively to the performance of her own firm. On the contrary, the overall package must have dynamic properties comparable to those of aggregate consumption. It is natural to identify the A_t^j component of the manager's remuneration with the "pay-for-luck" phenomenon identified in the empirical literature.[18] Our first-best contract thus implies that if the economy is doing well while an individual firm is doing badly, the manager of this particular firm may in fact see an increase in her overall compensation. It is not necessarily an abuse of the system if a well-compensated manager sees her overall compensation package increase even when her own firm is faltering. Far from being a sign of managerial entrenchment or inappropriate influence over the pay determination process (as in, e.g., Bebchuck and Fried (2003)), the presence of the "pay-for-luck" term induces the manager to select the correct intertemporal investment plan. We also note that, in production based asset pricing models such as the one considered here, aggregate dividends (and output) are very highly correlated with the return on the market portfolio. Our results are thus fully consistent with empirical work positively linking managerial compensation with aggregate stock market returns.

[18]This result is robust to a separation of workers and shareholders into distinct groups, with the latter being viewed as rentiers and the former as individuals who consumes their wages. In this case the optimal contract no longer retains the wage bill as an element of the "pay-for-luck' component. But "pay-for-luck" is still present in the form of the $d_t^{\tilde{j}}$ term.

What about one sided RPE? Let us assume, formally, that managers are of zero measure and appeal to Theorem 4. A modest adaptation of the argument of Section 4 to the multi-firm circumstances of this section reveals that the optimal contract must be of the form

$$g^j(d_t) = (\varphi \hat{d}_t^j + \varphi w_t n_t + \varphi d_t^{\tilde{j}})^\theta, \tag{28}$$

where $\theta = \frac{1-\gamma_s}{1-\gamma_m}$. If $\gamma_s > \gamma_m$, so that if shareholders are more risk-averse than managers, a not unlikely scenario, one-sided RPE will be observed as a natural consequence of optimal contracting since $g^j(\dot{})$ is a convex function. That is, the sensitivity of the managers' contract to aggregate state variables is smaller on the downside than it is on the upside. Only if the contract has this form will managers use the correct SDF.

In closing this section, let us observe that our contract specifies the same contract parameters, φ and ξ, for all firms. This is unlikely to be the case in reality. First, the implicit coordination necessary for firms to offer identical compensation contracts would constitute employment collusion and likely be illegal. A second, more relevant, reason is that across-firm differences in monitoring regimes or in the severity of firm-specific incentive problems may make the condition $\varphi\mu = 1$ unnecessary from the perspective of aligning the micro incentives of the manager with those of the shareholder-worker. In the same vein, if stock holding is not uniform across the population (incomplete markets in the form of limited stock market participation), the two elements of the manager's remuneration should not be weighted as per the aggregate NIPA income shares, but rather tailored to the distribution of wage income relative to capital income in effect for the firm's average shareholder (which may differ from firm to firm). Finally, one frequently observes firms offering contracts that are not linear but convex. In Danthine et al. (2008) we show that the incentives provided by convex contracts may, in some but not all

circumstances, well approximate the incentives provided by linear ones and partially make up for sub-optimal characteristics regarding the salary element of a manager's compensation package.

7. Bond trading

To what extent can active bond trading between the manager and the shareholder- workers make up for the absence of an optimal contract? We note at the outset that the information asymmetry between the manager and the consumer-worker-investors should preclude the manager trading in the firm's stock.[19] Bond trading is more plausible and the literature gives us strong reasons to suspect that it alone will promote a close convergence of the stochastic discount factors of the two agent types. Indeed the work of Heaton and Lucas (1996), Telmer (1995) and others demonstrates in various heterogenous agent contexts, that bond trading alone may be a close substitute for market completeness. The key issue therefore is whether the extent of bond trading necessary to promote the first best, or a close enough approximation thereof, is plausible in magnitude so as to be implementable in a market environment.

For simplicity let us again return to the setting of Section 4 ($\mu = 0$ and no effort decision). Since the manager must be given a contract of some sort - otherwise his decision problem is not well defined - the question can only meaningfully be posed in the context of some non-optimal specification. We explore the situation where $A_t \equiv 0$ (and the performance based remuneration is linear in free-cash-flow), and compare (1) allocations in otherwise identical delegated management economies with and without bondholding, and (2) bond trading allocations relative to the first best which we know can be achieved with $A_t + \varphi \hat{d}_t$ where $A_t \equiv \varphi w_t n_t$.[20] Simply stated, we seek to determine if bond trading can serve as a substitute for the absence of an optimally

[19] And indeed trading restrictions for stock market insiders are systematic.

[20] In the companion paper, we show that a manager's private wealth or a comfortable fixed remuneration both help in aligning SDFs although they are not full substitutes for a first-best contract.

designed, variable salary component in the manager's contract. Since the issue is a quantitative one, we resort to numerical computation. Appendix B describes the essential elements of our computational approach.

By the $\mu = 0$ assumption, the manager is effectively allowed to trade as many bonds as he wishes at the zero net supply price established by the shareholder-workers. By eliminating direct price reactions to changes in the manager's bond position, large quantity trades are unhindered in any way.[21] Under these modeling choices, the manager's decision problem can be represented as

$$\max_{\left\{n_t, i_t, b_t^m\right\}} E\sum_{t=0}^{\infty}\beta^t\left[u\left(c_t^m\right)\right] \tag{29}$$

s.t.

$$c_t^m = g^m\left(d_t\right) + b_{t-1}^m - q_t^b b_t^m,$$

$$d_t = f(k_t, n_t)\lambda_t - n_t w_t - i_t,$$

$$k_{t+1} = \left(1-\Omega\right)k_t + i_t,$$

$$c_t^m, i_t, n_t \geq 0,\ k_0,\ b_0^m \text{ given},$$

$$\lambda_{t+1} \sim dF\left(\lambda_{t+1};\lambda_t\right);\{w_t\},\{q_t^b\} \text{ given.}$$

where b_t^m is the manager's period t holdings of one period discount bonds and q_t^b is the period t bond price.

The corresponding problem of the consumer-worker-shareholder is

$$\max_{\{n_t^S, z_{t+1}, b_t^S\}} E\sum_{t=0}^{\infty}\beta^t[u(c_t^s) - \hat{H}(n_t^s)] \tag{30}$$

[21] By this statement we mean that we eschew any direct price consequences of the manager's bond demand. Indirectly, however, changes in investment or labor demand brought on by the initiation of bond trading will ultimately influence bond price behavior.

$$s.t. \quad c_t^s + q_t z_{t+1} + b_{t+1}^s q_t^b \leq \left(q_t + d_t\right) z_t + b_t^s + w_t n_t^s,$$

$$c_t^s, z_t, n_t^s \geq 0, \forall t;$$

$$\hat{s}_{t+1} \sim dG\left(\hat{s}_{t+1}; \hat{s}_t\right), \hat{s}_0 \text{given.}$$

In equilibrium market clearing requires:

$$\mu b_t^m + b_t^s \equiv 0$$

$$z_t \equiv 1$$

$$n_t^s = n_t^f \equiv n_t.$$

Table 1: Business Cycle Characteristics:
Admission of Bond Trading

	1. Optimal Contract, first-best economy		2. No bond trading, $A_t \equiv 0$		3. No bond trading, $A_t \equiv 0$ $\bar{b}^m = .01$		4. Bond trading, $A_t \equiv 0$, $\bar{b}^m = .01$	
	(a) [(a)]	(b)	(a)	(b)	(a)	(b)	(a)	(b)
y	1.80	1.00	1.01	1.00	1.01	1.00	1.80	1.00
c^m	.22	.87	.12	.16	.12	.16	.22	.39
d	8.08	-.97	.13	.16	.13	.16	8.08	-.97
c^s	.52	.87	.87	1.00	.87	1.00	.52	.87
i	5.74	.99	1.42	1.00	1.42	1.00	5.75	.99
k	.49	.35	.13	.26	.13	.26	.50	.35
n	1.37	.98	.14	.99	.14	.99	1.37	.98
w	.52	.87	.87	1.00	.87	1.00	.52	.87
r	.06	.96	.035	.99	.035	.99	.06	.96
b^m							3.05	-.30
q^b							.06	.96

[(a)] (a) Standard deviations in percent (b) $\text{Corr}(x, y)$

H-P filtered series. Parameters common to all cases: $\gamma_m = \gamma_s = 1, \alpha = .36, \beta = .99, \Omega = .025, \varphi = .01,$ $g^m(\ ) = \varphi d_t = \varphi \hat{d}_t$, since $\mu = 0$.

Table 1 summarizes our results. Panel 1 describes the quantitative properties of the benchmark economy under the first best contract. Standard deviations and correlations of the main

aggregates with output are provided. Panel 2 reports the properties of the same economy under the suboptimal contract with no bond trading permitted. Panel 2 underlines the major impact of suboptimal contracting: when endowed with a suboptimal contract, managers make employment and investment decisions that are hugely at variance with the first best decisions, resulting in an economy that is quantitatively very different from its first best counterpart (and noticeably displays very dampened fluctuations). Welfare evaluations follow accordingly. Panel 3 displays the properties of the same economy with suboptimal contracting and no bond trading but where managers are endowed with a small amount of initial wealth in the form of a non-tradeable bond position. The economy's statistical characteristics are unchanged from that of Panel 2: the small increase in wealth thus provided does not, per se, significantly alter the decisions of the manager.[22]

Panel 4 presents the results when bond trading is permitted. The striking feature to notice is its extraordinary effectiveness: the first best outcome is nearly perfectly replicated. As expected, trading the risk free asset without constraints in this economy with a single aggregate shock permits an almost perfect alignment of the stochastic discount factors of the two agent types. Nevertheless, the usefulness of bond trading as a device for aligning manager and shareholder-workers marginal rates of substitution is compromised by the enormous negative and positive bond positions that the manager will repeatedly assume along many feasible paths after a finite number of time periods.

[Insert Figure 2 here]

[Insert Figure 3 here]

[22]We include Panel 3 for the following reason: our computation methodology requires the manager's steady state bond holdings to be strictly positive. When contrasted with Panel 4 to follow, Panel 3 insures that the dramatic effects observed therein are not due simply to a wealth effect on the part of the manager.

This is seen in Figure 2 which presents a representative time path of managerial bond holdings relative to steady state consumption for the precise parametrization underlying Case 4 of Table 1. For the identical parametrization yet where the initial bond holdings is ten times smaller $b_0^m = \bar{b}^m = .001$, Figure 3 provides another observed time path.[23, 24] If, as in modern economies, wages are unable to collateralize loans, in either of these cases the manager rapidly becomes bankrupt. From these results it appears that security trading is not a plausible remedy for a suboptimal contract. If the contracting framework does not lead to a full alignment of the firm-owners and managers' SDFs, it is inevitable that the employment and investment decisions of the latter will deviate from those favored by the capital owners.

8. Related literature

We position our work in regard to the executive compensation, "pay-for-luck," dynamic equilibrium and related literatures.

8.1. Executive compensation: "Pay-for-Luck"

Bertrand and Mullainathan (2003) were the first to identify the "pay-for-luck" puzzle empirically. For the historical period they explore, CEO compensation is positively related to oil price shocks. Pay-for-luck style phenomena have subsequently been confirmed by many authors for various data periods (e. g., Garvey and Milbourne (2003)). One-sided RPE was first identified in Garvey and Milbourne (2006).

A large literature has ensued to explain these phenomena, and we present a review of some

[23] These samples are representative in the sense of being regularly observed. In other words, the probability of observing similar episodes over a long enough period of time is arbitrarily close to one.

[24] The enormous trading variation we see in these positions is not an artifact of the $\mu = 0$ assumption. In fact, we find that managerial bond trading is nearly as volatile when $\mu = .2$. This is the result of (1) relatively small equilibrium price variation and (2) the fact that the manager's bond position must vary roughly 5 times as much as the shareholder-worker's because his measure is so much smaller.

of the principal contributions below. In general, these explanations have taken the approach of rationalizing the dependence of a CEO's compensation on aggregate economy-wide variables. None of these explanations, however, emphasize the importance of the contract form in aligning the stochastic discount factors of shareholders workers and managers. Most are unable to accommodate risk-averse principals and agents, and none is set in the standard dynamic general equilibrium macro-finance context of the present study. In Goplan et al. (2007), CEOs choose the firm's strategy, and pay-for-luck induces them to provide effort to forecast sectoral returns and to adapt the firm's optimal response accordingly. In this model an absence of "pay-for-luck" causes the CEOs' investment decisions to be insensitive to sectoral events which is suboptimal since project returns depend on these co-movements. Empirical support for their conclusion comes from the fact that "pay-for-luck" is more typically observed in high technology, high R&D industries where strategic flexibility and innovation are of paramount importance.

Celentani and Loveira (2006) postulate a two firm scenario were the returns to effort for either manager are directly increasing in the aggregate state, and where the superior effort of either manager increases the likelihood of a favorable state ensuing. If the marginal benefits to managerial effort are sufficiently great in the favorable aggregate state, the increase in its likelihood implied by one manager's superior effort increases the likelihood of the other manager having exerted superior effort as well. An optimal contract thus rewards the manager when observables (aggregate output) are consistent with high effort one-sided RPE.

Himmelberg and Hubbard (2000) and Oyer (2004) emphasize CEO compensation as an equilibrium outcome in the labor market for scarce CEO talent. Oyer (2004) reasonably hypothesizes that a CEO's outside opportunities increase with increases in other firms' market values. Compensation schemes which do not filter out aggregate market movements thus serve to

index CEO compensation to outside market conditions thereby helping automatically to satisfy the CEO's participation constraint. In the same spirit Himmelberg and Hubbard (2000) analyze a model in which aggregate shocks -the bread and butter of macroeconomic fluctuations -increase the value of CEO services to firms. In labor markets where CEO services are inelastically supplied, the result of favorable aggregate shocks is a bidding-up of the value of all compensation packages, in particular those of the more talented, scarcer, CEOs, in a manner that appears to mimic "pay-for-luck." They provide empirical support for their assertion in the form of CEO compensation data which strongly suggests that the absence of RPE is limited to large, complex, business- cycle- sensitive firms. It is with respect to these firms that the supply of qualified CEO is limited. For small firms, where the supply of CEOs is relatively more elastic, however, compensation packages more typically reflect RPE.

Garvey and Milbourn (2003) provides a different motivation for RPE: RPE compensation effectively insures the CEO against systematic risk. Ceteris paribus, it declines when the market returns are high and rises when they are low. Its presence or absence in actual contracts may therefore reflect the extent to which CEOs rely on the firm to provide such insurance. Their empirical results suggest that wealthy CEOs, in general, place little value on such insurance, suggesting that these individuals can successfully manage overall market exposure by adjusting their own wealth portfolios. For the least wealthy class of managers they consider, however, RPE is much more the norm removing, on average, 80% of their market risk exposure.[25]

8.2. Dynamic Equilibrium

[25]This brief literature review is not exhaustive, and we acknowledge many other important contributions. Jenter and Kanaan (2006), for example, demonstrate the failure of RPE from the perspective of CEO dismissals: CEOs are more likely to be fired from their job after poor industry or poor market performance. In effect Boards of Directors do not filter out exogenous negative shocks from CEO retention decisions. Other authors explore the influence of noncompetitive product markets on the future on the structure of CEO compensation. Sala-Fumas (1992) and Aggrawal and Samwick (1999) are cases in point. We do not discuss these perspectives in full detail as our model, by construction, is not relevant for the issues they address.

Our emphasis has been to consider the implication of delegated management directly within the standard neoclassical paradigm. Under the optimal contract, the dynamics of our delegated management economy reproduce the stylized facts of the business cycle and we view this feature as one of the model's strengths. Aside from the basic informational asymmetry, the structure of the economy is entirely conventional. In contrast, one particular hypothesis prevalent in the corporate finance literature asserts that managers are "empire builders" (Jensen (1986)) who tend to over-invest and over-hire rather than return cash to the shareholders. A small number of studies have sought to explore its implications within the dynamic equilibrium paradigm. In Philippon (2006), managers have an inherent preference for size (firms with capital stock and labor resources exceeding their profit maximizing levels). Shareholders are more willing to tolerate such excesses in good times, a fact that tends to amplify the effects of aggregate shocks. In Dow et al. (2005), managers also prefer to maximize firm size. Their propensity to invest all available firm resources is held in check by the arrival, in the subsequent period, of costly auditors with the power to sequester excessive output. Albuquerque and Wang (2007) hypothesize a group of controlling shareholders (effectively acting as managers) who pursue private benefits by diverting resources from the firm. Such diversions are held in check by investor protections which vary across countries in their strength and effectiveness. Consistent with empirical regularities, their model demonstrates that countries with weaker investor protections should display overinvestment, larger risk premia etc. As made clear, we eschew these empire building or corrupt manager class of models, and postulate only that managerial preferences are defined over their own private consumption streams in a manner consistent with standard axiomatic foundations.

8.3. Other related work

A much larger literature has been concerned with optimal contracting between investors

and firm managers in the context of static one period partial equilibrium models. Analyzing as it does a wide range of principal-agent relationships, this literature is too large to be summarized here. In effect it has been concentrating on the performance-based element of a manager's remuneration and as such is somewhat orthogonal to the issue we have sought to confront. See Bolton and Dewatripont (2005) for a masterful review of this literature.

A final segment of literature attempts to rationalize the growing magnitude of executive compensation in particular as a multiple of worker compensations. Explanations run the gamet from rent extraction facilitated by enhanced managerial entrenchment (e.g., Bertrand and Mullainathan (2003)) to the demand for top talent which is better able to manage a larger resource base (Gabaix and Landier (2006); see this same reference for an excellent survey of the literature). Our decision to ignore the participation constraint of the manager constrains us from commenting on these issues.

The "quiet life" hypothesis, first expounded by Smith and Stulz (1985), focuses as we do on managers' risk appetite. It observes that, because they may be unable to diversify risk specific to their claims on the corporation, risk averse managers may display a tendency to forgo positive net present value projects. As observed in Panel 2 of Table 1, an inappropriately designed contract may indeed lead managers to adopt an excessively timid investment policy. Our analysis also shows that the problem is more related to the business cycle properties of free-cash-flows and the intertemporal elasticity of substitution of managers than it is to their risk position per se (see Danthine and Donaldson, 2008, for a detailed analysis). Bertrand and Mullainathan (2003) provides recent support for an enlarged definition of the quiet life hypothesis, one that includes a desire for peaceful labor relations in addition to a calmer-than-optimal investment policy.

9. Conclusions

In this paper we have studied the dynamic general equilibrium of an economy where risk averse shareholders delegate the management of the firm to risk averse managers. Our economy is one where information is asymmetric- the manager is better informed than shareholders. In some versions of our model, a moral hazard problem may be present as well when the non-observable effort of the manager is an input in production. We have derived the properties of the manager's optimal contract. This contract attains the first best and it results in an observational equivalence between the delegated management economy and the standard representative agent business cycle model.

The optimal contract has two main components: an incentive component that is proportional to free-cash-flow and is akin to a non-tradable equity position in the firm. And a variable "salary" component that is indexed to the aggregate wage bill and to aggregate dividends and may need to be corrected by a negative term proportional to aggregate consumption.

In our general equilibrium context it is thus not sufficient to resolve the "micro" level agency issues raised by delegation. Giving a share of dividends to self-interested managers with private information is an important requirement. Depending on the nature of the moral hazard and information problems, the share of free cash flows allocated to managers indeed may be very high. Yet, a simple minded application of this principle leads to endowing the manager with the wrong incentives. Because of the income and risk position she thus inherits, the manager will view the risks facing the firm through a lens - her own stochastic factor - that will possibly be widely at variance with the lens firm owners would like her to use. Aligning the stochastic discount factor is an essential component of the incentive problem. This second objective delineates the properties of the state dependent salary component. In our economy with a representative shareholder-worker, it is a linear function of the aggregate wage bill, the aggregate dividend and

aggregate consumption. As a byproduct of this contract form we are also able to evaluate the "pay-for-luck" and "one-sided RPE" phenomena as perfectly appropriate features of optimal contracting.

Does the absence of full information on the private wealth of the manager and on her market actions (including savings) detract from our message? This is a difficult question that has eluded the literature. We note the existence of two conflicting views. Some authors argue that the fact that managers are privately wealthy implies that they should be almost risk neutral at the margin. Although shareholders are supposed to be well diversified, they are not, however, risk neutral. Therefore if managers behave as if they are, they are not taking the decisions shareholders would want them to take. At the opposite extreme, the quiet life hypothesis argues that, compared to shareholders, managers are excessively invested in their own firm and thus insufficiently diversified. This, it is argued, suggests that they are likely to be excessively prudent, a fact that may justify convex performance based contracts. Our interpretation is that these two incompatible views reflect the fact that the principal-agent literature is ill-at-ease with the main lesson from asset pricing: the stochastic discount factor matters. In this paper we confront this difficult issue head on. The fact that managers may take private actions should not lead us to conclude that the stochastic discount factors of shareholders and managers will automatically be aligned. And the purported size of managers' income and wealth cannot mean that monetary incentives do not matter (or else the whole incentive debate is misguided!). If they do, it must be that even if managers' consumption is not tightly constrained by their compensation package, the latter indicates to them how the principals want them to view the world and in which light they should make the firm relevant decisions.

References

Aggarwal, R., Samwick, A., 1999. Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence, Journal of Finance 54, 1999-2043.

Albuquerque, R., Wang, N. Agency conflicts, investment and asset pricing, Journal of Finance, forthcoming.

Benveniste, L., Scheinkman, J., 1979. On the differentiability of the value function in dynamic models of economics, Econometrica 47, 727-732.

Bebchuk, L., Fried, J., 2004. Pay without Performance: The Unfulfilled Promise of Executive Compensation. Harvard University Press, Cambridge, MA.

Bertrand, M., Mullainathan, S., 2003. Enjoying the quiet life? Corporate governance and managerial preferences, Journal of Political Economy 111, 1043-1075.

Blackwell, D., 1965. Discounted dynamic programming, Annals of Mathematical Statistics 36, 226-235.

Bolton, P., Dewatripont, M., 2005. Contract Theory. MIT Press, Cambridge Mass.

Brock, W.A., 1982. Asset prices in a production economy.In: McCall, J.J. (Ed.), The Economics of Information and Uncertainty. University of Chicago Press, Chicago.

Celentani, M., Loveira, R., 2006. A simple explanation of the relative performance puzzle, Review of Economic Dynamics 9, 525-540.

Danthine, J.P., Donaldson, J.B., Gershun, N. 2008. Executive compensation: Salary vs. Incentive Pay: An inconvenient truth, working paper, Columbia University.

Donaldson, J.B., Gershun, N., Giannoni, M., 2010. Some unpleasant general equilibrium implications of executive incentive compensation contracts, working paper, Columbia University.

Dow J., Krishnamurthy, A., Gorton G., 2005. Equilibrium investment and asset prices under imperfect corporate control, American Economic Review 95, 659-681.

Fama, E.F., Jensen, M. C., 1983. Separation of ownership and control, Journal of Law and Economics 26 (2).

Gabaix, X. , Landier, A. , 2006. Why has CEO pay increased so much?, mimeo, Princeton University.

Garvey, G., T. Milbourn, 2003. Incentive compensation when executives can hedge the market: Evidence of relative performance in the cross section, The Journal of Finance 58, 1557-1581.

Garvey, G., T. Milbourn, 2006. Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad, Journal of Financial Economics 82, 197-225.

Goplan, R., Milbourn, T. , Song, F., 2007. Strategic flexibility and the optimality of pay for luck, mimeo, Olin Business School, Washington University in St. Louis.

Heaton, J., Lucas, D.J., 1996. Evaluating the effects of incomplete markets on risk sharing and asset pricing, The Journal of Political Economy 104, 443-487.

Himmelberg, C., Hubbard, R.G., 2000. Incentive pay and the market for CEOs: An analysis of pay for performance sensitivity, mimeo, Columbia Business School.

Jensen, M., 1986. The agency costs of free cash flow: Corporate finance and takeovers, American Economic Review 76, 323-330.

Jensen, M.C., Meckling, W.H., 2002. Specific and general knowledge, and organizational structure. In: Werin, L., Wijkander, H. (Eds), Main Currents in Contract Economics. Blackwell Press, Oxford.

Jenter, D., Kanaan, F., 2006. CEO turnover and relative performance evaluation, mimeo, MIT School of Management.

Lazear, E., 1995. Personnel Economics. MIT Press, Cambridge, Mass.

Lucas, R.E. Jr., 1978. Asset pricing in an exchange economy, Econometrica 46, 1429-1444.

Oyer, P., 2004. Why do firms use incentives that have no incentive effects? Journal of Finance 59, 1619-1649.

Philippon, T., 2006. Corporate governance over the business cycle, Journal of Economics Dynamics and Control 30, 2117-2141.

Prescott, E.C., Mehra, R., 1980. Recursive competitive equilibrium: The case of homogeneous households, Econometrica 48, 1365-1379.

Raith, M., 2005. Specific knowledge and performance measurement, mimeo, U. of Rochester.

Salas Fumas, V., 1992. Relative performance evaluation of management, International Journal of Industrial Organization 10, 473-489.

Shorish, S., Spear, S., 2005. Shaking the tree: An agency-theoretic model of asset prices, Annals of Finance 1, 51-72.

Smith, C.W., Stulz, R.M., 1985. The determinants of firms' hedging policies, Journal of Financial and Quantitative Analysis 20, 391-405.

Stein, J., 2003. Agency, information, and corporate investment. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier, 111-165.

Telmer, C., 1993. Asset pricing puzzle and incomplete markets, Journal of Finance 48, 1803-1832.

Appendix A

Proof of Theorem 3

With $\mu = 0$, the necessary and sufficient first order conditions for the 1st-best allocation of resources to the shareholder-workers are:

$$(c_t^s)^{-\gamma_s} \; f_1\left(k_t, n_t^s\right)\lambda_t = \hat{H}_1\left(n_t^s\right) \tag{A.1}$$

$$(c_t^s)^{-\gamma_s} = \beta \int \left(c_{t+1}^s\right)^{-\gamma_s} \left[f_1\left(k_{t+1}, n_{t+1}^s\right)\lambda_{t+1} + \left(1-\Omega\right)\right] dF\left(\lambda_{t+1};\, \lambda_t\right) \tag{A.2}$$

where $(c_t^s) = y_t - i_t$.

In the contracting economy with a competitive labor market, the necessary and sufficient first-order conditions governing labor supply, labor demand, and investment are

$$(c_t^s)^{-\gamma_s} \; w_t = \hat{H}_1\left(n_t^s\right) \tag{A.3}$$

$$w_t = f_1\left(k_t, n_t\right)\lambda_t \tag{A.4}$$

$$(c_t^m)^{-\gamma_m} \; g_1\left(x_t, \hat{s}_t\right) = \beta \int \left(c_t^m\right)^{-\gamma_m} \; g_1^m\left(x_{t+1}, \hat{s}_{t+1}\right)$$

$$\bullet \left[f_1\left(k_{t+1}, n_{t+1}^s\right)\lambda_{t+1} + \left(1-\Omega\right)\right] dF\left(\lambda_{t+1};\, \lambda_t\right) \tag{A.5}$$

In equilibrium, $n_t = n_t^s$, so that (A.3) and (A.4) together guarantee (23). Using the assumed contract forms, and recognizing that $\mu = 0$ implies $d_t = \hat{d}_t$, the manager's contract reduces to

$$g^m\left(x_t, \hat{s}_t\right) = \left(\varphi w_t n_t + \varphi d_t\right)^\theta = \left(\varphi\left(y_t - i_t\right)\right)^\theta . \tag{A.6}$$

Accordingly, equation (A.5) becomes

$$\left(\left(\varphi\left(y_t - i_t\right)\right)^\theta\right)^{-\gamma_m} \left(\varphi\left(y_{t+1} - i_{t+1}\right)\right)^{\theta-1} = \int \left(\left(\varphi\left(y_{t+1} - i_{t+1}\right)\right)^\theta\right)^{-\gamma_m} \bullet \left(\varphi\left(y_{t+1} - i_{t+1}\right)\right)^{\theta-1}$$

$$\bullet \left[f_1\left(k_{t+1}, n_{t+1}^s\right)\lambda_{t+1} + \left(1-\Omega\right)\right] dF\left(\lambda_{t+1};\, \lambda_t\right) \tag{A.7}$$

Equivalently,

$$\left(y_t - i_t\right)^{\theta(1-\gamma_m)-1} = \beta \int \left(y_t - i_t\right)^{\theta(1-\gamma_m)-1} \bullet \left[f_1\left(k_{t+1}, n_{t+1}\right)\lambda_{t+1} + \left(1-\Omega\right)\right] dF\left(\lambda_{t+1};\, \lambda_t\right) \qquad \text{(A.8)}$$

For equations (A.2) and (A.8) to coincide,

$$\theta\left(1-\gamma_m\right) - 1 = -\gamma_s \qquad \text{(A.9)}$$

If $\gamma_m \neq 1$ and $\gamma_s \neq 1$, equation (A.9) implies

$$\theta = \frac{1-\gamma_s}{1-\gamma_m} \qquad \text{(A.10)}$$

If $\left|\gamma_s - 1\right| > \left|\gamma_m - 1\right|$, and either $\gamma_s > 1$ and $\gamma_m > 1$ or $\gamma_s < 1$ and $\gamma_m < 1$, the contract is convex. If $\gamma_s = \gamma_m = 1$, any convexity parameter is optimal.

Proof of Theorem 4

To prove our result we assume that the contract has the stated form and show that the equilibrium conditions of the decentralized economy then coincide with the optimality conditions of the planner's problem.

Problem (1) of the shareholder is unchanged. The corresponding FOC's continue to be (13) and (14).

Under the proposed contract conditions (and taking note that, with full information, the relevant aggregate state variables are, again, $s_t = (k_t, \lambda_t)$), the manager's problem has recursive representation:[26]

[26]It again follows from Blackwell's (1965) Theorem and the results in Benveniste and Scheinkman (1979) that a continous, bounded $V^m(\ )$ exists that solves (A.2) provided $u(\cdot)$ and $f(\cdot)$ are increasing, continuous and bounded, and that $g^m(\cdot)$ is itself continuous and that $dF(A', \lambda'; A, \lambda)$ is also continuous with the property that for any continuous $h(k', A', \lambda')$, $\int h(k', A', \lambda') dF(A', \lambda'; A, \lambda)$ is also continuous in k and λ. In order for (A.3) and (A.4) to characterize the unique solution, the differentiability of $u(\cdot)$, $g^m(\cdot)$ and $f(\cdot)$ is required and

$$V^m\left(k_t,\lambda_t\right)=\max_{\{e_t,i_t,n_t\}} u\left(\varphi\hat{d}_t+A_t\right)-H(e_t)$$

$$+\beta\int V^m((1-\Omega)k_t+i_t,\lambda_{t+1})dF(\cdot), \tag{A.11}$$

The necessary and sufficient first order conditions to problem (A.2) can be written[27]

$$u_1\left(c_t^m\right)\varphi\left[f_2(k_t,n_t,\mu e_t)\lambda_t-w_t\right]=0, \tag{A.12}$$

$$-\varphi u_1(c_t^m)+\beta\int V_1^m(k_{t+1},\lambda_{t+1})dF=0, \text{where} \tag{A.13}$$

$$V_1(k_t,\lambda_t)=\varphi u_1(c_t^m)[f_1(k_t,n_t,\mu e_t)+(1-\Omega)] \tag{A.14}$$

$$u_1\left(c_t^m\right)\mu\varphi f_3(k_t,n_t,\mu e_t)\lambda_t=H_1(e_t). \tag{A.15}$$

Market clearing conditions (18), (19) and (20) apply. Equations (A.13) and (A.14) together imply that condition (7) is satisfied. Similarly, equation (A.12) together with (13) implies condition (6). Equation (A.15) reduces to (11) if and only if $\mu\varphi=1$.

Finally, for condition (8) to hold given (10) , one must have

$$\varphi\hat{d}_t+A_t=\varphi(y_t-w_tn_t-i_t)+A_t=\Delta(y_t-i_t),$$

for some scalar Δ. That is,

$$A_t=(\Delta-\varphi)(y_t-i_t)+\varphi w_tn_t\equiv\xi(y_t-i_t)+\varphi w_tn_t,$$

with $\xi=\Delta-\varphi$.

That $\xi\ll 0$ follows from the fact that the condition $\mu\varphi=1$ implies that the two first elements of the first best contract $\varphi\hat{d}_t+\varphi w_tn_t$ exhaust total output. There would be no value added remaining with which to compensate workers in the absence of the extra correction term.

$u(g^m(\cdot))$ must be concave. The assumptions made in this and the preceding footnote are maintained throughout the paper.

[27] In the usual spirit of a representative competitive firm the firm's manager is assumed not take account of the impact of her effort on the A_t term of her remuneration.

Proof of Theorem 5

As before, we postulate the form of the optimal contract and show that indeed this contract implements the first-best allocation. Under the postulated contract the representative manager j solves

$$V^j(k_0^j, \lambda_0^j, A_0^j; w_t) = \max_{\{c_t^j, e_t^j, n_t^j, i_t^j\}} E\sum_{t=0}^{\infty} \beta^t [u(c_t^j) - H(e_t^j)] \tag{A.16}$$

s.t.

$$c_t^j = g^j(d_t^j, s_t) = \varphi \hat{d}_t^j + A_t^j$$

$$d_t^j = f(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j - n_t^j w_t - \mu g^j(d_t^j, s_t) - i_t^j$$

$$k_{t+1}^j = (1-\Omega)k_t^j + i_t^j; k_0^j \text{given.}$$

$$c_t^j, d_t^j, i_t^j, n_t^j, e_t^j \geq 0$$

$$(s_{t+1}, \lambda_{t+1}^j) \sim dF(s_{t+1}, \lambda_{t+1}^j; s_t^j, \lambda_t^j)$$

Worker-shareholders are perfectly diversified. They collectively hold the market and are thus entitled to the aggregate dividend that we continue to identify as d_t. They consume the unique consumption good and equally share their working time n_t^s across all firms. Under these assumptions, problem (2) still perfectly represents the problem of the representative worker-shareholder. In particular condition (13) still holds.

The market clearing conditions are (19) and

$$\sum_{j=1}^{J} n_t^j = n_t^s$$

$$\sum_{j=1}^{J} i_t^j = i_t$$

$$\sum_{j=1}^{J} f(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j = y_t = c_t^s + \mu \sum_{j=1}^{J} c_t^j + i_t \tag{A.17}$$

Problem (A.16) yields the following conditions applying to all firms $j = 1,.., J$:

$$w_t = f_2(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j \tag{A.18}$$

which, in conjunction with(13), results in

$$\hat{H}_1(n_t^s) = u_1(c_t^s) f_2(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j.$$

This is the optimality condition (24). Optimal investment is determined by

$$1 = \beta \int \frac{u_1(c_{t+1}^j)}{u_1(c_t^j)} \left[f_1(k_{t+1}^j, n_{t+1}^j, \mu e_{t+1}^j)\lambda_{t+1}^j + (1-\Omega) \right] dF^j(.), \tag{A.19}$$

which is nothing but optimality condition (26).

The level of effort is given by the condition

$$u_1(c_t^j) f_3(k_t^j, n_t^j, \mu e_t^j)\lambda_t^j \mu\varphi = H_1(e_t^j), \tag{A.20}$$

from which one sees that

$$\mu\varphi = 1$$

is required to obtain the first best condition (27)[28].

Finally, we have to show that the Pareto Optimal risk sharing condition (27) is satisfied in equilibrium. To that end, let us first observe that the consumption of shareholders is proportional to ($y_t - i_t$). From the definition of the managers' contract, we have

$$c_t^j = \varphi\left[\hat{d}_t^j + w_t n_t + d_t^{\tilde{j}} \right] + \xi(y_t - i_t) \tag{A.21}$$

$$= \varphi\left[d_t + w_t n_t \right] + \mu\varphi c_t^j + \xi(y_t - i_t) \tag{A.22}$$

$$0 = \varphi c_t^s + \xi(y_t - i_t), \tag{A.23}$$

[28]This result implies that it would not be possible for managers to receive a firm-specific share of their firm's free-cash-flow

from which one obtains

$$c_t^s = -\frac{\xi}{\varphi}(y_t - i_t) = -\mu\xi(y_t - i_t).$$

Our second step is to note that the goods market clearing condition (A.17) implies that if the consumption of the shareholders is proportional to $(y_t - i_t)$, then the total consumption of management is as well:

$$\mu\sum_{j=1}^{J} c_t^j = y_t - i_t - c_t^s$$

$$= (1-\mu\xi)(y_t - i_t).$$

The last step consists of observing all managers' consumption levels are identical. This directly follows from

$$c_t^j = \varphi\left[\hat{d}_t^j + w_t n_t + d_t^{\tilde{j}}\right] + \xi(y_t - i_t)$$

$$= \varphi\left[d_t + w_t n_t\right] + \mu\varphi c_t^j + \xi(y_t - i_t)$$

$$= \varphi[y_t - i_t - \mu\sum_{j=1}^{J} c_t^j + \mu c_t^j] + \xi(y_t - i_t)$$

$$= (\varphi+\xi)(y_t - i_t) - \mu\sum_{i\neq j}^{J} c_t^i,\ j = 1,..,J \qquad \text{(A.24)}$$

Taking any arbitrary pair of equations in (A.24), say the kth and lth such equations and substract one from the other, one obtains

$$c_t^k(1-\mu) + (\mu-1)c_t^l = 0,$$

from which it is clear that $c_t^k = c_t^l$ and, as a consequence,

$$c_t^m = c_t^j = \frac{1}{1+\mu J}(\varphi+\xi)(y_t - i_t),\ \text{for}\ j = 1,2,..,J.$$

Appendix B

We used the standard log-linearization procedure for solving dynamic stochastic general equilibrium models suggested by Harald Uhlig (http://www2.wiwi.hu-berlin.de/institute/wpol/html/toolkit/version4_1.html). It requires choosing functional forms and parameter values and linearly approximating the equilibrium characterization in the neighborhood of the model's certainty steady-state. Model generated time series are first logged and then subjected to the Hodrick-Prescott filter, with the indicated quantities computed from these latter series. In all simulations $f(k_t, n_t)\lambda_t = k_t^{\alpha} n_t^{1-\alpha} \tilde{\lambda}_t$ where $\tilde{\lambda}_t$ is governed by $\tilde{\lambda}_{t+1} = \rho\lambda_t + \tilde{\varepsilon}_t$ with $\rho = .95$, $\sigma_{\varepsilon}^2 = .00712$, and $\alpha = .36$. As for preferences, $u(c_t^s) = \log c_t^s, u(c_t^m) = \log c_t^m, \hat{H}(n_t^s) = -Bn_t^s$ with $B = 2.85$ (see Hansen (1985) for a justification). Lastly, $\beta = .99, \Omega = .025, \varphi = .1$, where the contract form is $g^m(\hat{d}_t) = g^m(d_t) = \varphi d_t$ (recall that $\mu = 0$). Initial bond holdings, which equal steady-state bond holdings, are chosen from $\{.01, .001\}$.