Large Shareholders, Monitoring, and Ownership Dynamics: Toward Pure Managerial Firms?[1]

Amal Hilli
University Aix-Marseille

Didier Laussel
University Aix-Marseille

Ngo Van Long
McGill University

26th October 2012

[1] Address for correspondence: N. V. Long, department of economics, McGill University, 855 Sherbrooke St. West, Montreal, Quebec, H3A 2T7, Canada. Telephone: 1-514-389-4400 (ext. 00309). Fax: 1-514-398-4938. E-mail: ngo.long@mcgill.ca.

Abstract

We study ownership dynamics when the manager and the large shareholder, both risk neutral, simultaneously choose effort and monitoring level respectively to serve their non-congruent interests.We show that there is a wedge between the valuation of shares by atomistic shareholders and the large shareholder's valuation. At the Markov-perfect equilibrium, the large shareholder divests her shares. If the incongruence of their interests is mild, divestment is drastic: all her shares are sold immediately. If their interests diverge sharply, the divestment is gradual in order to prevent a sharp fall in share price. In the limit the firm becomes purely managerial.
JEL classification: G3
Keywords: Ownership dynamics; managerial firms.

1 Introduction

This paper develops a dynamic model of divestment by a large shareholder of a firm where her interest and that of the manager are not perfectly congruent. In our model, all shareholders, large and small, are risk neutral and have perfectly congruent objectives; however only the large shareholder monitors the manager while the small investors free ride on her monitoring effort.[1] We show how the degree of divergence of interests between the manager and the large shareholder affects the process of divestment. We demonstrate that when their interests diverge sharply, the divestment is gradual in order to prevent a sharp fall in share price. In the limit the firm becomes purely managerial, with a diverse ownnership, and no monitoring by shareholders. This paper thus serves to highlight a mechanism that lies behind the tendency for corporate governance to move gradually from concentrated to dispersed ownership, a pattern that has been observed over more than a century in major capitalist economies (such as Great Britain and the USA), and also more recently in countries such as Brazil. The key to our explanation is that the large shareholder cannot resist the temptation to sell shares when small investors' marginal benefit flow is greater than her own. While reducing her ownership (which entails a decrease in her monitoring effort) adversely affects the dividend flow to all investors, it does elicit more effort from the manager.

Berles and Means (1932) pointed to the transition to dispersed ownership in the US. Recent empirical work confirms this tendency.[2] For the U.K., the same tendency was reported in Scott (1990), and Franks et al. (2005), among others. Gorga (2009) documented a similar trend in Brazil from 1997 to 2002. Various reasons have been offered to explain the tendency for reduced concentration of ownership. Subrahmanyam and Titman (1999) argue that it becomes advantageous for firms to have a more dispersed ownership when

[1] These assumptions are also made in the seminal contribution of Burkart, Gromb, and Panunzi (1997), who however do not consider the dynamic process of divestment by the large shareholder. We build our dynamic model using the key elements of their static model.

[2] See, e.g., Demsetz and Lehn (1985), Mikkelson et al. (1997), LaPorta et al. (1999).

informational asymmetries between insiders and external investors are less important. Roe (1994) and LaPorta et al. (1999) attribute the dispersion of ownership in the US to the specific US laws and policies that discourage onwership concentration.

In this paper, we explain the tendency toward dispersed ownership by modelling, on the one hand, the trade-off between the gains from monitoring by a large shareholder and those from managerial initiatives, and on the other hand, the incentives for the large shareholder to divest (gradually, in typical cases) when her marginal valuation of ownership is below the small investors' valuation of the dividend stream that would arise on the assumption that she does not divest. The former aspect was investigated in an elegant static model by Burkart, Gromb, and Panunzi (1997). The latter aspect is built on the literature concerning the Coase conjecture. Coase (1972) argued that when a monopolist producing a durable good at constant marginal cost cannot commit, rational expectations by potential buyers, and his ability to sell repeatedly, would result in only one possible equilibrium outcome: he can only charge the price that would prevail under perfect competition, and the market demand is satisfied instantaneously.[3] In our model, where the large shareholder corresponds to the Coasian monopolist, we show that Coase's conjecture holds if the divergence of interests between the large shareholder and the manager is mild; in constrast, if this divergence is very strong, the Coase conjecture fails, and the large shareholder will divest only gradually, with share price falling slowly over time, converging only in the long run to the competitive price. There is also an intermediate case, in which at first the large shareholder undertakes a massive sale of shares, to be followed by a slow process of divestment of the remaining shares.

The intuition behind our results is simple. In all cases, the divestment is caused by the fact that small shareholders perceive that, under the assump-

[3] Coase's conjecture was confirmed by Stokey (1981), Gul et al. (1986), and others, under the assumptions that marginal cost is constant and the interval between successive sales shrinks to zero. Coase's conjecture fails if there is increasing marginal cost (Kahn, 1987), or depreciation (Karp, 1996).

tion that the large shareholder would not divest, their dividend stream per share is worth more than the large shareholder's marginal returns on a share (as she has to incur the monitoring cost). This wedge in marginal valuations implies that equilibrium must involve share trading. When the divergence of interests between the manager and the large shareholder is mild, her total instantaneous payoff (net of monitoring cost) is a strictly concave and increasing function of her fraction of ownership. Therefore the revenue she would obtain from selling her shares at the competitive share price strictly dominates the present value of the stream of her instantaneous payoff obtained from maintaining her initial stock. Hence her optimal policy is to sell off all her shares in one go. In the reverse case, the strong divergence of interests implies that her total instantaneous payoff is a strictly convex and increasing function of her fraction of ownership. The equilibrium share price function must in this case equal the large shareholder's capitalised marginal instantaneous payoff, which increases in her shareholding. Selling shares too quickly would cause a drastic fall in share price. So it is optimal for her to sell gradually.[4]

Our paper is related to a strand of literature which deals with the dynamic process of adjustment of shareholding based on the insight from the literature on the Coase conjecture. Unlike our model specification which places emphasis on the conflict between the manager and the large shareholder, Gomes (2000) assumes that the large shareholder is also the manager of the firm. In that model, the owner-manager is playing a share-selling game against the collection of small investors. The gains from trade arise because by selling her shares, the owner-manager can diversify idiosyncratic risks with investors. The investors perceive that the owner-manager may be of one type or another. Although the owner-manager knows her type, investors know only the probability distribution of types. At each period, the owner-manager moves

[4] In the intermediate case, which involves weak congruence of interests, coupled with a large stake, the large shareholder's instantaneous payoff function is S shaped: it is convex (concave) when her fraction of ownership is small (large). Then the large shareholder's optimal strategy is to make an initial lumpy sale of a fraction of her shares, to be followed by a time path of gradual sale of the remaining fraction.

by choosing her new fraction of equity ownership and her effort level which is unobservable. Investors update their belief about the owner-manager's type, and they price shares in the market accordingly. Gomes shows that when outside investors face this adverse selection problem, the owner-manager's equilibrium strategy involves divesting her shares gradually over time (in contrast to the perfect information benchmark, where the owner-manager would sell all her shares in the first period). This gradualism is necessary for the entrepreneur to develop a reputation for treating minority shareholders well.

Gomes's conclusion that a risk-averse owner-manager would divest shares gradually over time is re-enforced by DeMarzo and Urošević (2006) who show (in a model with moral hazard instead of adverse selection) that if moral hazard is weak enough, the large shareholder trades immediately to the competitive price-taking allocation. With strong moral hazard, however, she will adjust her stake gradually. DeMarzo and Urošević (2006) emphasize the large shareholder's trade-off between risk diversification (which calls for a small shareholding) and her incentives and ability to improve the firm's performance (which increases with her fraction of ownership of her firm). DeMarzo and Urošević assume that the utility function of the large shareholder exhibits constant absolute risk aversion. Her wealth consists of a risk-free account and risky shares in her firm. Her sale strategy is motivated by consumption smoothing and risk diversification. When she sells her shares, investors anticipate a decrease in her effort. Hence, when reducing her stake, she is likely to generate a decrease in share price.[5] Edelstein et al. (2007) generalize the model of DeMarzo and Urošević (2006) to a setting with multiple strategic insiders. They show that the aggregate stake of the insiders decreases gradually over time, and that the long run equilibrium aggregate stake of the insiders are greater for firms with a larger number of insiders.

[5] The authors also noted that the time-inconsistency problem, raised by Coase, also applies to their model. They consider both the benchmark case where the large shareholder can commit ex ante to an ownership policy, and the case where commitment is not possible.

In contrast, in our model, all agents (the large shareholder, the small investors, and the manager) are risk neutral: their utility function is linear in income. Moreover, we focus on the divergence of interests between the large shareholder and the manager (who does not own shares): in each period, there is an agency problem occuring between the large shareholder and the manager.[6] The separation of management (by the manager) and control (by the large shareholder) is a major driving force behind the dynamics of share sales. By divesting, the large shareholder can influence the time path of the equilibrium effort level of the manager, as well as the time path of her own level of monitoring of his action. In our model, the wedge between her marginal valuation of a share and that of the atomistic investor indeed arises from a strategic effect, namely Nash equilibrium managerial effort (in the game between the manager and the large shareholder) being decreasing in the fraction of shares held by the large shareholder. This strategic effect is absent in DeMarzo and Urošević (2006). While DeMarzo and Urošević (2006) obtained the result that total divestment is the ultimate outcome if control benefits are small relative to risk aversion, the mechanisms driving the dynamic process in the two models are quite different.

The paper is organized as follows. Section 2 describes the basic framework, drawn from Burkart, Gromb, and Panunzi (1997). Section 3 deals with the commitment benchmark. Section 4 turns to time-consistent strategies and characterizes the Markov-perfect equilibrium corresponding to different regions of the parameter space. Section 5 discusses the main contributions of our model in relation to the literature on divestment. Section 6 concludes.

2 The Model

Our model is a dynamic extension of the static model proposed by Burkart, Gromb, and Panunzi (1997), or BGP for short.

[6] Our setting is based on the seminal paper of Burkart, Gromb, and Panunzi (1997), who restricted attention to a static setting. Our dynamic analysis has added some interesting features about the equilibrium divestment strategies, as outlined above.

2.1 The basic static setting

Consider a firm run by a manager, denoted by M. He manages the firms and owns no shares. There is a large shareholder, denoted by L. She owns a fraction α of shares but does not manage the firm. She has a strong incentive to monitor the manager when α is large. The remaining fraction $1-\alpha$ is owned by a continuum of atomistic shareholders who free ride on the monitoring effort of L. All agents are risk neutral.

In each period, the firm must choose one project to carry out. The projects it faces come in four known types, denoted by $i = \{0, 1, 2, 3\}$. There are two mutually exclusive states of nature, denoted by A and B. A type i project, if carried out, will yield a pair of benefits (Π_s^i, b_s^i) in state s $(s = A, B)$ where Π_s^i is verifiable and accrues to the shareholders, while b_s^i is non-verifiable and accrues to the manager. Assume that, for both $s = A$ and $s = B$, $\Pi_s^0 = b_s^0 = 0$ and that both Π_s^1 and b_s^1 are large negative numbers, say $\Pi_s^1 = b_s^1 = -k_s\Pi$ where $\Pi > 0$ and k_s is a large positive number. The pair of benefits associated with a type 2 project is $(\Pi, b) \gg (0, 0)$ if the state of nature is A, and is $(\Pi, 0)$ if the state of nature is B. Each type 3 project yields the pair (Π, b) if the state of nature is A, and the pair $(0, b)$ if the state of nature is B.[7] State A occurs with probability $\lambda < 1$. The state of nature occurs before the firm chooses its project. However information about which state has occured is not revealed to the manager unless he exercises effort, in which case he will obtain the information with some probability. The numbers Π, b, k_s and λ are common knowledge. The properties of type 2 and type 3 projects are summarized in Table 1.

PLEASE PLACE TABLE 1 HERE

For simplicity, assume that each period the firm is presented with exactly four projects, one from each type. Assume that everyone knows which of these

[7]BGP's specification of the payoffs is slightly different from ours. However the analysis of the static model is essentially the same for both specifications.

four projects is a type 0 project. The remaining three projects, however, are presented to the firm as named projects (γ, μ, θ). Everyone knows that in period t there is a one-to-one mapping ϕ_t from the set $\{\gamma, \mu, \theta\}$ to the set of project types $\{1, 2, 3\}$. However, if the manager does not spend some effort $e > 0$ in information-seeking activities, this mapping will not be revealed to the manager (nor to the shareholders).

If M exercises effort level e in period t, where $e \in [0, 1]$, then with probability e, he will be completely informed, i.e., he will discover both (i) the mapping ϕ_t and (ii) whether the state of nature in period t is A or B. By investing e in the information-seeking activities, then, with probability e, all the uncertainty is eliminated for the manager, but with probability $1 - e$, he will remain completely uninformed.

The large shareholder, L, does not observe the manager's choice of effort level e. She chooses her monitoring effort level E, where $E \in [0, 1]$, to attempt to find out the information that the manager has obtained. If M remains completely uninformed, then L learns that M knows nothing. If M is completely informed, then L will find out that M is informed, and with probability E she captures all of his information (about the mapping ϕ_t and the state of nature) while with probability $1 - E$ she obtains no information. It is assumed that neither e nor E is verifiable, and that M and L must make their choice (e, E) simultaneously.

Notice that the parameter λ is a measure of the congruence of interests between the manager and the shareholders. In the polar case where $\lambda = 1$, there would be perfect congruence of interests. In what follows, we assume that $0 < \lambda < 1$.

After the choices (e, E) have been made, there are three possible cases. First, if both parties remain uninformed, they will rationally agree that the firm should choose the type 0 project.[8] Second, if M is the only informed party, then L knows that M will assure himself the payoff b (since he knows

[8] Recall that k is a large positive number, therefore the expected payoff from a random choice of projects is negative.

both the true mapping ϕ_t and the state of nature), which implies that, from L's vantage point, the income accruing to the shareholders is Π with probability λ and zero with probability $1-\lambda$. Third, if both parties are informed, L, knowing which of (γ, μ, θ) is the type 2 project, will exercise her control right and require M to undertake the type 2 project. Carrying out the type 2 project implies that M's payoff is b with probability λ and 0 with probability $1-\lambda$.

Let us reproduce below BGP's derivation of the Nash equilibrium pair (e, E).

M's effort cost is $(1/2)e^2$. Given E, he chooses $e \in [0,1]$ to maximize his expected payoff

$$-\frac{1}{2}e^2 + (1-e) \times 0 + e \times \{E\left[\lambda b + (1-\lambda) \times 0\right] + (1-E)b\}$$

The first order condition yields M's downward-sloping reaction function:

$$e = \min\{1, b\left[1-(1-\lambda)E\right]\}. \tag{1}$$

This implies that the large shareholder's monitoring will reduce the manager's incentives to exercise effort. Taking into account the fact that $E \in [0,1]$ and assuming that $0 < b < 1$, we deduce that manager's chosen effort level is in the interior of his range of feasible choices $[0,1]$.

Denote by $D(e, E)$ the expected aggregate dividends for the shareholders. Then

$$D(e,E) \equiv \{(1-e) \times 0 + e \times \left[E\Pi + (1-E)\lambda\Pi\right]\} = e\Pi\left[\lambda + (1-\lambda)E\right] \tag{2}$$

Note that for a fixed E, an increase in the manager's effort raises the expected dividends

$$\frac{\partial D}{\partial e} = \Pi(\lambda + (1-\lambda)E) > 0 \tag{3}$$

Similarly, for a fixed $e > 0$, an increase in the shareholder's monitoring effort raises the expected dividends:

$$\frac{\partial D}{\partial E} = e(1-\lambda)\Pi > 0. \tag{4}$$

Assume that L's effort cost is $(1/2)E^2$. Given e, the large shareholder will choose $E \in [0,1]$ to maximize her expected payoff,

$$-\frac{1}{2}E^2 + \alpha D(e, E) \tag{5}$$

Her first order condition at an interior solution is

$$\alpha D_E = E \iff \alpha(1-\lambda)e\Pi = E \tag{6}$$

i.e. the large shareholder equates her marginal cost of monitoring to the marginal increase in her expected dividend income that results from increased monitoring.

Thus we obtain L's upward-sloping best-reply function:

$$E = \min\{1, \alpha\Pi(1-\lambda)e\} \tag{7}$$

The following assumption is made.

Assumption A: $0 < \lambda < 1$, $b < 1$, and

$$b\Pi < \frac{1}{\lambda(1-\lambda)}.$$

This is a technical assumption to ensure that the Nash equilibrium pair (e, E) is in the interior of the feasible set $[0,1] \times [0,1]$. Indeed, the condition $0 < b < 1$ ensures that the manager's reaction function $e(E)$ has the property that $e(0) = b < 1$ and $e(1) = \lambda b < 1$. On the other hand, the large shareholder's reaction function $E(e)$ has the property that $E(0) = 0$ and $E(\lambda b) = \alpha(1-\lambda)\lambda b\Pi < 1$. Thus the two reaction curves must intersect each other inside the unit square, $[0,1] \times [0,1]$. Allowing equilibrium at a corner would not affect our qualitative results, but would make the analysis more unwieldy.

It follows from Assumption A and equations (1) and (7) that the Nash equilibrium satisfies

$$E(\alpha) = \frac{\alpha\Pi b(1-\lambda)}{1+\alpha\Pi b(1-\lambda)^2} < 1 \text{ and } e(\alpha) = \frac{b}{1+\alpha\Pi b(1-\lambda)^2} < 1 \tag{8}$$

Let us define $\Omega \equiv \Pi b$. The parameter Ω may be regarded as an index of the "stake" of the game between the manager and the large shareholder. Given α, the Nash equilibrium monitoring effort $E(\alpha)$ depends on the two parameters Ω and λ.[9] Note that

$$E'(\alpha) = \frac{\Omega(1-\lambda)}{\left[1+\alpha\Omega(1-\lambda)^2\right]^2} > 0 \text{ and } e'(\alpha) = \frac{-b\Omega(1-\lambda)^2}{\left[1+\alpha\Omega(1-\lambda)^2\right]^2} < 0. \quad (9)$$

Thus the Nash equilibrium effort level e of the manager is decreasing in the large shareholder's fraction of ownership, and the latter's monitoring effort, E, increases with her fraction of ownership.

Denote the expected aggregate dividends (or equity value) by $W(\alpha) \equiv D(e(\alpha), E(\alpha)) = e(\alpha)\Pi\left[\lambda + (1-\lambda)E(\alpha)\right]$. From (8) and (2),

$$W(\alpha) \equiv D(e(\alpha), E(\alpha)) = \frac{\Omega\left[\lambda + \Omega\alpha(1-\lambda)^2(1+\lambda)\right]}{\left(1+\Omega\alpha(1-\lambda)^2\right)^2}. \quad (10)$$

Equity value, $W(\alpha)$, is in general a non-monotone function of the ownership fraction α of the large shareholder. To see this formally, let us denote by F the set of all admissible couples (λ, Ω) that satisfies Assumption A:

$$F \equiv \left\{(\lambda, \Omega) \in \mathbb{R}^2_+ | \quad 0 < \Omega < \frac{1}{\lambda(1-\lambda)} \text{ and } 0 < \lambda < 1\right\} \quad (11)$$

The upper boundary of F is the U-shaped curve $\Omega = \frac{1}{\lambda(1-\lambda)} \equiv f(\lambda)$, where $\lim_{\lambda\to 0} f(\lambda) = \lim_{\lambda\to 1} f(\lambda) = \infty$, as depicted in Figure 1.

PLEASE PLACE FIGURE 1 HERE

Since

$$W'(\alpha) = \frac{\Omega^2(1-\lambda)^3\left(1-\alpha\Omega(1-\lambda^2)\right)}{\left(1+\Omega\alpha(1-\lambda)^2\right)^3} \quad (12)$$

[9] The greater is Ω, the greater is the equilibrium monitoring effort level of the large shareholder. However, the Nash equilibrium level of monitoring $E(\alpha; \Omega, \lambda)$ is not monotone in λ. For values of λ close to unity, a marginal increase λ (i.e. a higher degree in congruence on interests) leads to lower equilibrium monitoring.

we conclude that $W'(\alpha) > 0$ for all $\alpha \in (0,1)$ if and only if the admissible couple (λ, Ω) belongs to the set Q defined below:

$$Q \equiv \left\{ (\lambda, \Omega) \in \mathbb{R}^2_+ | \quad 0 < \Omega < \frac{1}{1-\lambda^2} \text{ and } 0 < \lambda < 1 \right\} \subset F. \tag{13}$$

Note that Q is a proper subset of F. The upper boundary of region Q is the dotted curve depicting the equation $\Omega = \frac{1}{1-\lambda^2} \equiv q(\lambda) < f(\lambda)$. Along this curve, as $\lambda \to 0$, $\Omega \to 1$, and as $\lambda \to 1, \Omega \to \infty$. Clearly, if $(\lambda, \Omega) \in Q$, then equity value $W(\alpha)$ is maximized at the corner $\alpha = 1$.

It is also easy to verify that if $(\lambda, \Omega) \in F - Q$ then there exists a unique value $\widehat{\alpha} \in (0,1)$, which depends on λ and Ω, such that (i) $W(\alpha)$ is maximized at $\alpha = \widehat{\alpha}$, where

$$\widehat{\alpha} \equiv \frac{1}{\Omega(1-\lambda^2)} < 1 \text{ for } (\lambda, \Omega) \in F - Q.$$

(ii) $W'(\alpha) > 0$ if $\alpha < \widehat{\alpha}$, and (iii) $W'(\alpha) < 0$ if $\alpha > \widehat{\alpha}$.

Define α_2^* to be the fraction of shares owned by the large shareholder that would maximize equity value:

$$\alpha_2^* \equiv \arg \max_{0 \leq \alpha \leq 1} W(\alpha)$$

Then

$$\alpha_2^* = \min \left\{ 1, \frac{1}{\Omega(1-\lambda^2)} \right\}.$$

Consider the large shareholder's net income (after subtracting her effort cost):

$$R(\alpha) \equiv \alpha W(\alpha) - \frac{1}{2} \left[E(\alpha) \right]^2. \tag{14}$$

It can be verified that $R'(\alpha)$ is strictly positive for all $\alpha \in (0,1)$:

$$R'(\alpha) = \frac{\Omega\lambda + \alpha\Omega^2(1-\lambda)^2(1+\lambda)}{(1+\alpha\Omega(1-\lambda)^2)^3} > 0 \text{ for all } \lambda \in (0,1). \tag{15}$$

The sum of the instantaneous payoffs to the large shareholder and the collection of small shareholders is called "net equity value," defined as

$$V(\alpha) = R(\alpha) + (1-\alpha)W(\alpha) = W(\alpha) - \frac{1}{2} \left[E(\alpha) \right]^2$$

BGP showed that net equity value is maximized at[10]

$$\alpha_1^* = \frac{1}{\frac{1}{1-\lambda} + \Omega(1-\lambda^2)} < 1. \tag{16}$$

Remark 1 BGP did not report an interesting fact, which we state below as fact 1.

Fact 1 The marginal value of a share to the large shareholder is smaller than its value to an atomistic shareholder. That is, $R'(\alpha) < W(\alpha)$.

The proof is as follows. From (14),

$$R'(\alpha) = W(\alpha) + \alpha W'(\alpha) - E(\alpha)E'(\alpha) \tag{17}$$

Therefore

$$R'(\alpha) - W(\alpha) = \alpha W'(\alpha) - E(\alpha)E'(\alpha) = \alpha D_e \frac{de(\alpha)}{d\alpha} < 0 \tag{18}$$

More explicitly, using (15) and (10),

$$R'(\alpha) - W(\alpha) = -\frac{\alpha\left(\lambda + (1+\lambda)\alpha\Omega(1-\lambda)^2\right)\Omega^2(1-\lambda)^2}{\left(1+\alpha\Omega(1-\lambda)^2\right)^3} \leq 0 \text{ for all } \alpha \in [0,1] \tag{19}$$

with equality only at $\alpha = 0$.The inequality $R'(\alpha) < W(\alpha)$ reflects the fact that the small shareholders are free-riding on the monitoring effort of the large shareholder, who incurs monitoring costs without being compensated. These costs are non-verifiable, and therefore the large shareholder cannot "bill" the firm for her monitoring effort.

Fact 1 suggests the following conjecture: if share trading is allowed to take place at each point of time, the only time-consistent equilibrium outcome is that the large shareholder will sell her shares, either in a lumpy fashion, or gradually, or both. This conjecture will be shown to be correct (see Section 4).

[10] See their Proposition 1. Clearly, $\alpha_1^* < \alpha_2^*$.

For the analysis of the dynamic adjustment process in the following sections, it is important to determine whether $R(\alpha)$ is strictly convex, or strictly concave, or neither. It turns out that this depends on the value taken by the couple (λ, Ω). The following Lemma is useful.

Lemma 1 *Consider the following subsets of F, denoted by X, T and A, where $X \cup A \cup T = F$,*

$$X \equiv \left\{ (\lambda, \Omega) \in \mathbb{R}_+^2 | \quad 0 < \lambda \leq \frac{1}{2} \text{ and } 0 < \Omega \leq \left(\frac{1-2\lambda}{2-2\lambda} \right) \frac{1}{1-\lambda^2} \right\} \tag{20}$$

$$T \equiv \left\{ (\lambda, \Omega) \in \mathbb{R}_+^2 | \quad \frac{1}{2} < \lambda < 1 \text{ and } 0 < \Omega < \frac{1}{\lambda - \lambda^2} \right\} \tag{21}$$

$$A \equiv \left\{ (\lambda, \Omega) \in \mathbb{R}_+^2 | \quad 0 < \lambda \leq \frac{1}{2} \text{ and } \left(\frac{1-2\lambda}{2-2\lambda} \right) \frac{1}{1-\lambda^2} < \Omega < \frac{1}{\lambda - \lambda^2} \right\} F \tag{22}$$

Then $R(\alpha)$ is (i) strictly convex in α for all $\alpha \in (0,1)$ if $(\lambda, \Omega) \in X$, (ii) strictly concave in α for all $\alpha \in (0,1)$ if $(\lambda, \Omega) \in T$, and (iii) is S-shaped (convex for all α in the open interval $(0, \widetilde{\alpha})$ and convex for all α in the open interval $(\widetilde{\alpha}, 1)$ if $(\lambda, \Omega) \in A$, where

$$\widetilde{\alpha} \equiv \frac{1-2\lambda}{2\Omega(1-\lambda)(1-\lambda^2)}. \tag{23}$$

Proof From (15),

$$R''(\alpha) = \frac{\Omega^2 (1-\lambda)^2 \left[(1-2\lambda) - 2\alpha\Omega(1-\lambda)^2(1+\lambda) \right]}{\left(1+\alpha\Omega(1-\lambda)^2\right)^4} \tag{24}$$

(i) If (λ, Ω) is in X, then

$$2\Omega \leq \frac{1-2\lambda}{(1-\lambda)(1-\lambda^2)}$$

implying that for all $\alpha < 1$, $2\Omega\alpha(1-\lambda)(1-\lambda^2) < 1-2\lambda$ hence $R'' > 0$.

(ii) If (λ, Ω) is in T, then $(1-2\lambda) < 0$, hence $R'' < 0$ for all $\alpha \in [0,1]$.

(iii) If (λ, Ω) is in A, $(1-2\lambda) - 2\alpha\Omega(1-\lambda)^2(1+\lambda)$ can be of either sign. Then define $\widetilde{\alpha}(\lambda, \Omega)$ by

$$0 < \widetilde{\alpha}(\lambda, \Omega) \equiv \frac{1-2\lambda}{2\Omega(1-\lambda)(1-\lambda^2)} < 1 \text{ for } (\lambda, \Omega) \in A,$$

we can see that $R''(\alpha) > 0$ for $0 < \alpha < \widetilde{\alpha}(\lambda, \Omega)$ and $R''(\alpha) < 0$ for $\widetilde{\alpha}(\lambda, \Omega) < \alpha < 1$. ■

Remark 2 The upper boundary of region X is the curve

$$\Omega = \frac{1-2\lambda}{2(1-\lambda)^2(1+\lambda)} = \left(\frac{1-2\lambda}{2-2\lambda}\right)\frac{1}{1-\lambda^2} \equiv h(\lambda) \text{ for } 0 < \lambda < \frac{1}{2}$$

Along this curve, as $\lambda \to 1/2, \Omega \to 0$. As $\lambda \to 0, \Omega \to 1/2$. This negatively-sloped curve lies below the curve $\Omega = \frac{1}{1-\lambda^2} \equiv q(\lambda)$.

3 A dynamic version of the model: the commitment benchmark

Let us now assume that the projects mentioned above last for only one period (or, more precisely, since we use continuous time, for an arbitrarily small time interval). Assume that at each instant t, a new set of projects become available. At t the manager exercises effort level $e(t)$ and the large shareholder chooses her monitoring level $E(t)$. If the large shareholder's ownership fraction at t is $\alpha(t)$, her equilibrium instantaneous payoff is her expected dividends minus her effort costs,

$$R(\alpha(t)) \equiv \alpha W(\alpha(t)) - \frac{1}{2}\left[E(\alpha(t))\right]^2. \tag{25}$$

Suppose the large shareholder contemplates reducing her fraction of ownership at the rate $-\dot{\alpha}(t)$ at time t. (We allow $\dot{\alpha}(t)$ to be of either sign.) Let $p(t)$ be the market price of a share at time t. Recall that the number of shares is fixed and normalized at unity. We assume that investors have rational expectations, so that the share price at t is simply the value of the discounted stream of expected dividends:

$$p(t) = \int_t^\infty \exp(-r(\tau - t))W(\alpha(\tau))d\tau. \tag{26}$$

where r is the interest rate. Differentiating (26) with respect to t yields

$$\dot{p}(t) = rp(t) - W(\alpha(\tau)). \tag{27}$$

This equation is the usual non-arbitrage condition in a competitive asset market: the return to holding an asset (i.e. the sum of capital gains and dividends) is just equal to the opportunity cost, $rp(t)$, of foregone interest income. The payoff to the large shareholder is then

$$J^c(\alpha_0) = \int_0^\infty \exp(-rt)\left[R(\alpha(t)) - \dot{\alpha}(t)p(t)\right]dt \tag{28}$$

where $-\dot{\alpha}(t)p(t)$ is flow of cash receipts generated by her divesting rate $-\dot{\alpha}(t)$, subject to $\alpha(0) = \alpha_0$ and $1 \geq \alpha(t) \geq 0$.

What is her optimal divesting strategy? The answer to this question depends on whether she can commit to a time path of sale of her shares. Let us begin with the the benchmark case of commitment.

Suppose that the large shareholder is able to commit to a whole time path of her shareholding $\alpha(t)$. Her objective function (28) can then be simplified as follows. Let us write

$$\phi(t) \equiv \exp(-rt)p(t) = \int_t^\infty \exp(-r\tau)W(\alpha(\tau))d\tau$$

Since W is bounded, it is clear that $\lim_{t\to\infty}\phi(t) = 0$. Then

$$\int_0^\infty \exp(-rt)p(t)\dot{\alpha}(t)dt = \int_0^\infty \phi(t)\dot{\alpha}(t)dt =$$

$$= [\alpha(\infty)\phi(\infty) - \alpha(0)\phi(0)] - \int_0^\infty \dot{\phi}(t)\alpha(t)dt =$$

$$-\alpha(0)\int_0^\infty \exp(-rt)W(\alpha(t))dt + \int_0^\infty \alpha(t)\exp(-rt)W(\alpha(t))dt$$

Thus her objective function becomes

$$\max_{0<\alpha(t)\leq 1}\int_0^\infty \exp(-rt)\left[R(\alpha(t)) + (\alpha_0 - \alpha(t))\,W(\alpha(t))\right]dt$$

Obviously, the solution is to choose the same value for $\alpha(t)$ for all $t \in (0,\infty)$. This is because the maximum path can be found by point-wise differentiation with respect to $\alpha(t)$ for each $t \in (0,\infty)$. The first order condition is identical for each t. Hence $\alpha(t)$ is the same for all $t > 0$.

In words, it is optimal to make an immediate jump in the state variable (an impulse control) to some optimal commitment level $\alpha(0^+) = \alpha^c$, and after this initial jump, $\alpha(t)$ will be kept constant at α^c for ever, where α^c is the value of α that maximizes $R(\alpha) - (\alpha - \alpha_0)W(\alpha)$ subject to $\alpha \in [0, 1]$. That is, the large shareholder chooses her immediate net acquisition, $\alpha(0^+) - \alpha_0$, to maximize the capitalised value of her time-invariant dividend flow (net of her effort cost), $R(\alpha)/r$, minus the cost of acquisition (i.e., the price of $W(\alpha)/r$ per share, multiplied by the number of shares acquired, $\alpha(0^+) - \alpha_0$). Note that in principle $\alpha(0^+) - \alpha_0$ can be positive (acquisition of shares) or negative (sale of shares). However, we will show below that the properties of the function $Z(\alpha; \alpha_0) \equiv R(\alpha) + (\alpha_0 - \alpha)\, W(\alpha)$ imply that the optimal $\alpha(0^+) - \alpha_0$ is negative (i.e. sales of shares).

We now show that $Z(\alpha; \alpha_0)$ is strictly quasi-concave in α, attaining its maximum at some $\alpha^c \in (0, \alpha_0)$. Using the definition of Z, we obtain the derivative

$$Z_\alpha(\alpha; \alpha_0) \equiv R'(\alpha) - [W(\alpha) - (\alpha_0 - \alpha)\, W'(\alpha)]$$

For ease of economic interpretation, divide both terms by r.The term $[W(\alpha) - (\alpha_0 - \alpha)W'(\alpha)]\, /r$ represents the marginal revenue from divesting, while the term $R'(\alpha)/r$ is her marginal cost of divesting (it is measured as the fall in the capitalised value of her dividend flow (net of her effort cost) caused by a marginal reduction in her share ownership). Let α^c be the value of α such that the two terms are equalized. Let us show that α^c is in the interior of the interval $[0, \alpha_0]$, and that α^c is an increasing function of the initial α_0. The proof is as follows.

From the definition of $R(\alpha)$, i.e.(25), we have the identity

$$R'(\alpha) \equiv \alpha W'(\alpha) + W(\alpha) - E(\alpha)E'(\alpha) \text{ for all } \alpha \in [0, 1]\,.$$

So the first order condition for maximizing $Z(\alpha; \alpha_0)$ reduces to

$$Z_\alpha(\alpha; \alpha_0) \equiv \alpha_0 W'(\alpha) - E(\alpha)E'(\alpha) = 0$$

This condition gives

$$\alpha^c = \frac{\alpha_0(1-\lambda)}{1+\alpha_0(1-\lambda)^2(1+\lambda)\Omega} \equiv \chi(\alpha_0)$$

$$= \frac{1}{\frac{1}{\alpha_0(1-\lambda)} + (1-\lambda)(1+\lambda)\Omega} \in (0,1)$$

It is easy to verify that second order condition is satisfied whenever the first order condition is satisfied, indicating that the maximum is unique.

Note that α^c is smaller than α_0 :

$$\frac{\alpha^c}{\alpha_0} = \frac{(1-\lambda)}{1+\alpha_0(1-\lambda)^2(1+\lambda)\Omega} = \frac{1}{\frac{1}{1-\lambda} + \alpha_0(1-\lambda)(1+\lambda)\Omega} < 1.$$

Proposition 1 (Optimal asset sale strategy under commitment). *If the large shareholder can make a binding commitment on her time path of share holding, her optimal policy is to reduce her shareholding immediately from her initial holding α_0 to a commitment level α^c where*

$$\alpha^c = \frac{\alpha_0(1-\lambda)}{1+\alpha_0(1-\lambda)^2(1+\lambda)\Omega} \equiv \chi(\alpha_0) < \alpha_0$$

and afterward she retains her remaining shares for ever.

Remark 3 The commitment level of asset holding, α^c, is generally smaller than the level α_1^*, which was reported by BGP as the level that maximizes net equity value; see our equation (16).[11] The two levels are equal only in the special case that $\alpha_0 = 1$.

The large shareholder is willing to commit not to sell more shares thereafter because she wants to elicit a higher initial share price. Rational buyers would be willing to pay this price only if they believe she would not go on reducing her shares and consequently reducing her monitoring effort level. However, rational buyers are smart. They know the formula offered in Proposition 1, and know that the next instant, the current α^c will become the new α_0 and therefore the large shareholder will sell again, attaining a new

[11] We thank a referee for pointing this out.

and lower α^c, and so on. In other words, the solution described in Proposition 1 displays the property of time-inconsistency. The commitment strategy of holding $\alpha(t) = \chi(\alpha_0)$ for all $t \in (0, \infty)$ implies that at any time $t_1 > 0$, if the large shareholder would be released from her original commitment, she would again want to sell immediately some more shares (because at t_1 the relevant initial holding would be α_{t_1}) and consequently reduce her monitoring effort. Then the share price would fall below the initial price, $W(\alpha^c)/r$. Her additional sales would inflict capital losses to the previous buyers of shares, because they have been fooled into believing that the large shareholder would sell assets only once and would maintain her commitment level of monitoring. Solutions that display time-inconsistency are generally regarded as unacceptable (Coase, 1972). Therefore we must look for time-consistent solutions.

4 Markov perfect equilibrium

In this section, we seek solutions that have the time-consistent property, and, in addition, that would be robust to perturbation. More precisely, we are insisting on a stronger property than time-consistency, namely Markov perfect equilibrium.[12] In a Markov perfect equilibrium, the large shareholder uses a Markovian strategy ω and the market has a Markovian price function, or expectation rule, ρ (which we will explain in more detail below) such that (i) given ρ, the Markovian strategy ω maximizes L's payoffs, for all possible starting (date, state) pairs (t, α_t), and (ii) given ω, the Markovian price function ρ is consistent with rational expectations.[13]

Assume that the atomistic agents all have a common Markovian price expectation function $p(t) = \rho(\alpha(t))$, where ρ is a function of the state variable α. The price expectations function must be rational, in the sense that the

[12] For an exposition of the concepts of time-consistency and Markov perfect equilibrium, and a proof that Markov perfect equilibria are time-consistent, see Dockner et al. (2000). Long (2010) provides some simple examples.

[13] For some examples of Markovian price function in the industrial organization literature, see Karp (1996), Driskill and McCafferty (2001), and Laussel et al. (2004).

share price must equal the capitalized value of the future dividend stream:

$$\rho(\alpha(t)) = \int_t^\infty \exp(-r(\tau - t))W(\alpha(\tau))d\tau, \tag{29}$$

where $\{\alpha(.)\}_t^\infty$ is the time path of the state variable α induced by the strategy ω of the large shareholder, from time t, when the state variable takes the value α_t.

A strategy ω of the large shareholder is a specification of (i) a collection of disjoint intervals $I_1, I_2, ..., I_m$ where $I_i \equiv [a_i, b_i] \subset [0,1]$, (ii) a lumpy sale function $L_i(.)$ that specifies a downward jump in the state variable, such that $\alpha - 1 \leq L_i(\alpha) \leq \alpha$, (if $L_i(\alpha)$ is negative, it signifies a lumpy purchase of shares), and (iii) a gradual sale function $g(.)$ defined for all $\alpha \notin I_i$, such that

$$\dot{\alpha}(t) = -g(\alpha(t)) \text{ for } \alpha \notin I_i$$

where $g(\alpha) \in (-\infty, \infty)$.

The payoff to the large shareholder, given (t, α_t), is

$$\int_t^\infty \exp(-r(\tau - t)) \left[R(\alpha(\tau)) - \dot{\alpha}(\tau)p(\tau)\right] d\tau$$

where $p(\tau) = \rho(\alpha(\tau))$.

Definition: *A Markov-perfect equilibrium is a pair (ρ, ω) such that, (i) given the price function ρ, the strategy ω maximizes the large shareholder's payoff, starting at any (date,state) pair (t, α_t), and (ii) given ω and (t, α_t), the price function ρ satisfies the rational expectation properties (29).*

Remark 4: Equation (29) yields the usual non-arbitrage condition (27).

4.1 Markov perfect equilibrium in the parameter region X (convex $R(\alpha)$)

By Lemma 1, in region X the large shareholder's instantaneous returns function $R(\alpha)$ is strictly convex for all $\alpha \in (0,1)$. Furthermore, since $W(0) = R'(0)$, it follows that $R(\alpha) > \alpha W(0)$ for all $\alpha > 0$. This means that, starting with α_0, if the large shareholder were to sell all her α_0 instantaneously, her share would be sold at the price $p = \frac{1}{r}W(0)$, and her payoff (revenue from sales) would be $\frac{\alpha_0}{r}W(0)$, which is strictly smaller than $\frac{1}{r}R(\alpha_0)$, her payoff if she does not offer to sell her shares. This suggests that selling her shares gradually would be better than selling them off in one go. The following proposition confirms this intuition.

Proposition 2: *If (λ, Ω) is in the set X (defined by (20)), so that the shareholder's net returns function $R(\alpha)$ is convex, then the large shareholder's equilibrium strategy is to sell her shares gradually, such that $\alpha(t) \to 0$ asymptotically as $t \to \infty$, and the atomistic investors' equilibrium price function is*

$$\rho(\alpha) = \frac{1}{r}\left[\frac{\Omega\left(\lambda + \alpha\Omega(1-\lambda)^2(1+\lambda)\right)}{\left(1+\alpha\Omega(1-\lambda)^2\right)^3}\right] = \frac{1}{r}R'(\alpha) \tag{30}$$

where the equilibrium price is increasing in the fraction of shares held by the large shareholder:

$$\rho'(\alpha) = \frac{1}{r}R''(\alpha) > 0 \text{ for all } \alpha \in (0,1) \text{ and for all } (\lambda, \Omega) \in X.$$

The large shareholder's optimal rate of sale at time t is $-\dot{\alpha}(t)$, where

$$\frac{-\dot{\alpha}(t)}{\alpha(t)} = \frac{\left(\lambda + \alpha\Omega(1-\lambda)^2(1+\lambda)\right)\Omega^2(1-\lambda)^2}{\left(1+\alpha\Omega(1-\lambda)^2\right)^3 N\rho'(\alpha)} > 0 \tag{31}$$

Along the path of of disinvestment, the share price falls monotonically, converging asymptotically to $\rho(0) = \frac{1}{r}W(0) = \frac{1}{r}R'(0)$.

Proof:

The Hamilton-Jacobi-Bellman (HJB) equation for the large shareholder is

$$rJ(\alpha) = \max_{\dot{\alpha}}\left\{R(\alpha) - \dot{\alpha}\rho(\alpha) + J'(\alpha)\dot{\alpha}\right\} \tag{32}$$

Since this equation is linear in $\dot{\alpha}$, the optimal $\dot{\alpha}$ is finite only if

$$\rho(\alpha) = J'(\alpha) \text{ for all } \alpha \in (0,1) \tag{33}$$

Substituting this into the HJB equation (32), we obtain

$$J(\alpha) = \frac{1}{r}R(\alpha) \text{ for all } \alpha \in (0,1)\,. \tag{34}$$

Thus, the value function evaluated at α is just equal to the discounted stream of net returns that would be obtained if α were kept constant for ever.[14]

Then the equilibrium price function is

$$\rho(\alpha) = J'(\alpha) = \frac{1}{r}R'(\alpha) \tag{35}$$

All the necessary conditions for an equilibrium are satisfied. Let us verify that this is indeed better than selling off all of α in one go. The latter action would yield a return of $\frac{1}{r}R(0) + \frac{1}{r}\alpha_0 W(0) = \frac{1}{r}\alpha_0 W(0)$.In region X, the function $R(\alpha)$ is strictly convex in α. This strict convexity and $R(0) = 0$ implies that $R(\alpha) > R'(0)\alpha$ for all $\alpha > 0$. But from (19), $R'(0) = W(0)$. Therefore $R(\alpha) > W(0)\alpha$ for all $\alpha > 0$.This shows that selling gradually is better than selling all α off in one go.

Finally let us characterize the selling strategy and the time path of sales. In view of equations (19) and (35), we have $\rho(\alpha) < W(\alpha)/r$. This inequality indicates that along the equilibrium divestment path, share price is lower than the (hypothetical) present value of a constant stream of dividend per share calculated on the hypothesis that α remains unchanged through time.[15]

Recall that $p(t) = \rho(\alpha(t))$. Since we require that the no-arbitrage condition (27) holds, we must have

$$\rho'(\alpha)\dot{\alpha} = r\rho(\alpha) - W(\alpha)$$

i.e.,

$$\rho'(\alpha)\dot{\alpha} = r\rho(\alpha) - W(\alpha) = R'(\alpha) - W(\alpha)$$
$$\dot{\alpha}(t) = -\frac{\alpha\left(\lambda + \alpha\Omega(1-\lambda)^2(1+\lambda)\right)\Omega^2(1-\lambda)^2}{\left(1+\alpha\Omega(1-\lambda)^2\right)^3\rho'(\alpha)}$$

[14] When the function $R(\alpha)$ is strictly convex, the large shareholders gains nothing by selling gradually as compared with keeping α for ever, but she must sell gradually in the Markov perfect equilibrium. For, if she instead held on to her α_0, the expected share price would be constant for ever at $W(\alpha_0)/r$, which would of course induce her to sell.

[15] Why is the price lower than the current level of dividend per share? The key to the answer lies in the fact that in region X, $W'(\alpha) > 0$.Therefore, as the large shareholder sells more and more shares, she exerts less and less monitoring effort, and the dividend per share falls. The investors know this. They would not pay a price equal to the present value of a constant stream of current dividend.

Using (24) to substitute for $\rho'(\alpha)$,

$$\frac{\dot{\alpha}}{\alpha} = \frac{r\left(\lambda + \alpha\Omega(1-\lambda)^2(1+\lambda)\Omega^2(1-\lambda)^2\right)\left(1+\alpha\Omega(1-\lambda)^2\right)}{\left[-1+2\lambda+2\alpha\Omega(1-\lambda)^2(1+\lambda)\right]\Omega^2(1-\lambda)^2} < 0 \text{ in region } X.$$

Note that as $\alpha \to 0$, $\dot{\alpha}/\alpha$ tends to a negative constant. Thus $\alpha(t) \to 0$ asymptotically.

We must show that given the price rule (30), no deviation from the trading strategy (31) can increase the large shareholder's expected discounted profits above $J(\alpha) = R(\alpha)/r$. Consider a given lumpy sale L_i deviation taking α to $\alpha - L_i$ where L_i is in the interior of the interval $(\alpha - 1, \alpha)$.[16] After the deviation, the large shareholder uses the equilibrium trading rule, the value of the continuation game is $J(\alpha - L_i)$, which implies, by (35), that $J'(\alpha - L_i) = \rho(\alpha - L_i)$. So the expected payoff obtained from this deviation is $J(\alpha - L_i) + L_i\rho(\alpha - L_i)$. The best among such interior deviations is found by differentiating $J(\alpha - L_i) + L_i\rho(\alpha - L_i)$ with respect to L_i. This gives the necessary condition for the best L_i is $-J'(\alpha-L_i)-L_i\rho'(\alpha-L_i)+\rho(\alpha-L_i) = 0$. Since $J'(\alpha - L_i) = \rho(\alpha - L_i)$, we have $J''(\alpha - L_i) = \rho'(\alpha - L_i)$, so the above necessary condition reduces to $L_i J''(\alpha - L_i) = 0$. But $J'' > 0$. It follows that $L_i^* = 0$, i.e. the best interior deviation is zero deviation. Let us now consider a jump to $\alpha = 1$. This deviation yields the payoff $\frac{1}{r}\left[R(1) - (1-\alpha)W(1)\right]$, while sticking to the candidate equilibrium strategy yields $\frac{1}{r}R(\alpha)$. Now the strict convexity of $R(\alpha)$ in region X implies that $R(\alpha) - R(1) > R'(1)(\alpha - 1)$. Thus $R(\alpha) > R(1)-(1-\alpha)R'(1) > R(1)-(1-\alpha)W(1)$, because $W(1) > R'(1)$ by (19). We conclude that there is no profitable discontinuous deviation.$\blacksquare$

To summarize, when parameter values are in region X, the equilibrium outcome is that the large shareholder sells her shares gradually. If she were to sell them off all in one go, her payoffs would be lower. Since the share price function $\rho(\alpha)$ is increasing in α, when the large shareholder divests, the price declines. If she were to divest all in one go, the share price would fall too sharply.

[16] The case of deviation that causes a jump to 0 has been examined above. Deviation to 1 is examined separately below.

Interestingly, the fraction of shares held by large shareholder never vanishes in finite time. Starting from any initial fraction α_0, the time it takes to reduce her holding to a given fraction $\underline{\alpha} > 0$ is increasing in λ and decreasing in r. Suppose for instance that she initially holds $\alpha = 1$, $\Omega = 0.2$ and $r = 0.05$. If $\lambda = 0.1$, the time it takes for α to falls to 0.1 is 29 years. If $\lambda = 0.2$, the corresponding time is 10.5 years. Figure 2 plots the time paths under these parameter values.

PLEASE PLACE FIGURE 2 HERE

4.2 Markov perfect equilibrium in the parameter region T (concave $R(\alpha)$)

In region T, the instantaneous returns function $R(\alpha)$ is strictly concave for all $\alpha \in (0, 1)$. Furthermore, since $W(0) = R'(0)$, it follows that $R(\alpha) < \alpha W(0)$ for all $\alpha > 0$. This means that, starting with α_0, if the large shareholder were to sell all her α_0 instantaneously, her share would be sold at the price $p = \frac{1}{r} W(0)$, and her payoff (revenue from sales) would be $\frac{1}{r} W(0)\alpha_0$, which is strictly greater than $\frac{1}{r} R(\alpha_0)$, which is her payoff if she does not offer to sell her shares. This suggests that selling her shares gradually would be worse than selling them all off in one go. The following result confirms this intuition.

Lemma 2 *When $R(\alpha)$ is concave, the policy of gradual sales/purchases cannot be an equilibrium.*

Proof Suppose we were to try the HJB equation (32) as in the preceding subsection, then, under the assumption of gradualism, we would have come up with the implication that $\rho(\alpha) = \frac{1}{r} R'(\alpha)$ and that $\dot{\alpha}/\alpha > 0$, because $\frac{1}{r} R''(\alpha) < 0$ in region T. This would imply that, given α_0, the large shareholder would purchase additional shares gradually. In particular, as $\alpha \to 1$, the equation would then reduce to

$$\frac{\dot{\alpha}}{\alpha} = \frac{r\left(\lambda + \Omega(1-\lambda)^2(1+\lambda)\Omega^2(1-\lambda)^2\right)\left(1 + \Omega(1-\lambda)^2\right)}{\left[-1 + 2\lambda + 2\Omega(1-\lambda)^2(1+\lambda)\right]\Omega^2(1-\lambda)^2}$$

Since $\lambda > 1/2$ in region T, the right hand side is a constant, implying that α will reach 1 at some finite t_1. We now show that the implied share price path $p(t)$ would not satisfy the rational expectations requirement (29). Indeed, for all $t < t_1$, we would have

$$p(t) = \rho(\alpha(t)) = \frac{1}{r}R'(\alpha(t))$$

Therefore

$$\lim_{t \uparrow t_1} p(t) = \frac{1}{r}R'(1).$$

But, from (29) and (19),

$$p(t_1) = \frac{1}{n}\int_t^{\infty} \exp(-r(t-\tau))W(1)d\tau = \frac{W(1)}{r} > \frac{1}{r}R'(1).$$

This would imply an upward jump in shares price at time t_1, which would not be consistent with rational expectations. (Atomistic investors, expecting such an upward jump, would refuse to sell their shares before t_1, defeating the large shareholder's gradual purchase scheme.)$\blacksquare$

Proposition 3 *When the parameter vector* (λ, Ω) *is in region* T*, the unique Markov perfect equilibrium consists of the price function* $\rho(\alpha) = \frac{1}{r}W(0)$ *for all* $\alpha \in [0,1]$ *and the lumpy sale strategy* $L(\alpha) = \alpha$ *for all* $\alpha \in [0,1]$*. The payoff to the large shareholder is*

$$J(\alpha) = \frac{\alpha W(0)}{r} > \frac{R(\alpha)}{r} \tag{36}$$

where the inequality is strict for all $\alpha \in (0,1]$*.*

Proof

(i) Given the lumpy sale strategy $L(\alpha) = \alpha$ (i.e., given that the large shareholder sells off all her shares at the initial instant), rational expectations imply that the atomistic traders must hold linear price function $\rho(\alpha) = \frac{1}{r}W(0)$.

(ii) Given the linear price $\rho(\alpha) = W(0)/r$, the large shareholder's problem is to maximize, given any $\alpha_0 \in [0,1]$,

$$J(\alpha_0) = \max \int_0^{\infty} e^{-rt}\left[R(\alpha(t)) - \dot{\alpha}(t)\rho(\alpha(t))\right] dt = \int_0^{\infty} e^{-rt}\left[R(\alpha(t)) - \frac{1}{r}\dot{\alpha}(t)W(0)\right] dt$$

Integration by parts, noting that $\alpha(t)$ is bounded, yields

$$\int_0^\infty \dot{\alpha}(t)\left[\frac{1}{-r}e^{-rt}\right]dt = \frac{\alpha(0)}{r} - \int_0^\infty \alpha(t)e^{-rt}dt = \int_0^\infty \left[\alpha_0 - \alpha(t)\right]e^{-rt}dt$$

Thus

$$J(\alpha_0) = \max \int_0^\infty e^{-rt}\left[R(\alpha(t)) + (\alpha_0 - \alpha(t))\,W(0)\right]dt$$

Since $R(\alpha) + (\alpha_0 - \alpha)W(0)$ is strictly concave in α when the parameter vector (λ, Ω) is in region T, the solution is trivially to set $R'(\alpha(t)) = W(0)$, i.e. $\alpha(t) = 0$ for all t. This means a downward jump in α at time zero.

It follows that $J(\alpha_0) = (\alpha_0/r)W(0)$. Since $R(\alpha)$ is strictly concave and $R'(0) = W(0)$ by equation (19), we obtain (36).

(iii) Gradual sales/purchases cannot be an equilibrium, as showed in Lemma 2.■

The intuition behind Proposition 3 is as follows. In Region T, the interests of the manager and the shareholders diverge only mildly. The overall tendency is to reduce the large shareholder's stake in the firm. Since $R(\alpha)$ is concave, if she were to sell her shares gradually, the required share price function would be decreasing in α, which would imply that it would pay to reduce α to zero as quickly as possible so as to get the highest possible price.

4.3 Markov perfect equilibrium when $R(\alpha)$ is S-shaped

When the vector of parameter (λ, Ω) is in region A, the instantaneous returns function $R(\alpha)$ is convex in the range $[0, \widetilde{\alpha}]$ and concave in the range $(\widetilde{\alpha}, 1]$ where $\widetilde{\alpha}$ is defined by (23).In this region, the interests of the manager and the shareholders diverge sharply, as in region X, but the absolute possible benefits for both parties are larger than in region X. From our analysis in the two preceding sub-sections, it becomes clear that the Markov perfect equilibrium in this case would be for the large shareholder to make an initial lumpy sale of part of her stock, if $\alpha_0 > \widetilde{\alpha}$. Once this asset position $\widetilde{\alpha}$ is reached, she will start a gradually sale policy, liquidating her shares asymptotically. We formalize this in Proposition 4.

Proposition 4 *When the vector of parameter* (λ, Ω) *is in region* A, *then if* $\alpha > \widetilde{\alpha}$ *the large shareholder will divest immediately in a lumpy fashion the fraction of her stock in excess of* $\widetilde{\alpha}$. *Afterwards, she gradually divests the remaining shares. The atomistic investors hold the following price expectation rule*

$$\rho(\alpha) = \begin{cases} \frac{1}{r}R'(\widetilde{\alpha}) & \text{if } \alpha \in [\widetilde{\alpha}, 1] \\ \frac{1}{r}R'(\alpha) & \text{if } \alpha \in [0, \widetilde{\alpha}] \end{cases}. \tag{37}$$

The value function of the large shareholder is

$$J(\alpha) = \begin{cases} \frac{1}{r}R(\widetilde{\alpha}) + (\alpha - \widetilde{\alpha})\frac{U'(\widetilde{\alpha})}{r} & \text{if } \alpha \in [\widetilde{\alpha}, 1] \\ \frac{1}{r}R(\alpha) & \text{if } \alpha \in \alpha \in [0, \widetilde{\alpha}] \end{cases}.$$

The lumpy sale function is $L(\alpha) = \alpha - \widetilde{\alpha}$ *for all* $\alpha \in [\widetilde{\alpha}, 1]$. *The gradual sale rule is*

$$g(\alpha) = -\dot{\alpha} = \frac{\alpha\left(\lambda + \alpha\Omega(1-\lambda)^2(1+\lambda)\right)\Omega^2(1-\lambda)^2}{\left(1+\alpha\Omega(1-\lambda)^2\right)^3 \rho'(\alpha)} > 0 \text{ for } \alpha \in \alpha \in [0, \widetilde{\alpha}] \tag{38}$$

Proof

(i) Given that $L(\alpha) = \alpha - \widetilde{\alpha}$ for all $\alpha \in [\widetilde{\alpha}, 1]$, rational expectations on the part of atomistic investors imply that $\rho(\alpha) = \rho(\widetilde{\alpha})$ for all $\alpha \in [\widetilde{\alpha}, 1]$. Given $g(\alpha)$ defined by (38), for all α in $[0, \widetilde{\alpha}]$, the same argument as that used in the proof of Proposition 2 applies to show that the price rule $\rho(\alpha) = \frac{1}{r}R'(\alpha)$ satisfies the rational expectations requirement.

(ii) Given (37), any deviation by the large shareholder inplying a discontinuous variation in α in the interval $[0, \widetilde{\alpha}]$ can be ruled out, as was shown in the proof of Proposition 2. Given that $\rho(\alpha)$ is constant in $[\widetilde{\alpha}, 1]$, any deviation implying a jump in α from one value to another value $\widehat{\alpha} \neq \widetilde{\alpha}$ is ruled out by an argument similar to that used in part (ii) of the proof of Proposition 3.

(iii) Given (37), any deviation by the large shareholder implying an upward jump from some $\alpha' < \widetilde{\alpha}$ to some $\alpha'' > \widetilde{\alpha}$ would yield a present value equal to $\frac{1}{r}\left[R(\alpha'') - R'(\widetilde{\alpha})(\alpha'' - \alpha')\right]$ whereas sticking to the candidate equilibrium yields $\frac{1}{r}R(\alpha')$. We can show that this is not profitable, because, from the convexity of R for $\alpha < \widetilde{\alpha}$, it holds that $\frac{1}{r}R(\alpha') > \frac{1}{r}\left[R(\widetilde{\alpha}) - R'(\widetilde{\alpha})(\widetilde{\alpha} - \alpha')\right]$, while from the concavity of R for $\alpha > \widetilde{\alpha}$, it holds that $\frac{1}{r}\left[R\left(\widetilde{\alpha}\right) - R'\left(\widetilde{\alpha}\right)\left(\widetilde{\alpha} - \alpha'\right)\right] > \frac{1}{r}\left[R'\left(\alpha''\right) - R'\left(\widetilde{\alpha}\right)\left(\alpha'' - \alpha'\right)\right]$.

Finally, consider a deviation that implies a downward jump from some $\alpha' > \widetilde{\alpha}$ to some $\alpha'' < \widetilde{\alpha}$. This would yield a value $J(\alpha'') + (\alpha' - \alpha'')J'(\alpha'')$, which, given the convexity of J for values of $\alpha \leq \widetilde{\alpha}$, is lower than the value $J(\widetilde{\alpha}) + (\alpha' - \widetilde{\alpha})J'(\widetilde{\alpha})$ obtained by following the equilibrium path. These arguments also rule out any other candidate equilibrium.■

5 Discussion

We have shown that even in a model without risk aversion and without the involvement of the large shareholder in managing the firm, there are non-trivial dynamics of share prices and divestment. This is in contrast to DeMarzo and Urošević (2006) who showed that "risk aversion or control benefits are necessary to create a wedge in the large shareholder's and investors' marginal benefit flow that will induce trade" (p. 785). On the other hand, in a reduced form of their general model, they obtained similar dynamics. The first task of this section is therefore to discuss the similarities and differences between the two reduced form models.

Both models rely on a wedge between the value to the large shareholder and the value to the small investors that drives trade in shares. In both models, when actions take place in continuous time, the reduction in the large shareholder's stake in the firm can be gradual or instantaneous, or a combination of the two. Which of the three cases hold depends, in both models, on the convexity or concavity of the function that describes share value to the large shareholder.

However, behind these similarities, there is a fundamental difference. DeMarzo and Urošević (2006) postulate that there there are control benefits to the large shareholder. In particular, they assume that the marginal control benefit, at any given effort level of the large shareholder, is generally non-zero. This marginal control benefit is depicted by the term $b(a, t)$ in their equation (4), p. 782. On the same page, $b(a, t)$ is defined as the reduction in the effort cost $f(e, a, t)$, holding effort constant,[17]

$$b(a,t) = -f_2(e(a,t), a, t)$$

They justify this marginal control benefit term by stating that "this specification generalizes the standard moral hazard setting by allowing the cost to depend on the holding a. It may be easier to monitor and impliment changes with a large stake in the the firm, or there may be other benefits associated with control" (p. 778). Without this marginal control benefit term, there will be no trade in their model when agents are risk neutral. The direction of trade depends on the relative strength of the marginal control benefit in relation to the coefficient of absolute risk aversion.

In our model, by assumption, the large shareholder has no private benefits of control, and all agents are risk neutral. So what is the mechanism that drives trade in our model? The answer is that there are strategic interactions between the non-owner manager and the non-managing large shareholder. At any ownership level of the large shareholder, there is a static game between her and the manager. The static Nash equilibrium effort levels of these two players turn out to depend on the size of their stakes. This strategic effect is absent in the model of DeMarzo and Urošević (2006). The strategic effect is the novel feature that allows non-trivial dynamics to exist in our model with risk-neutral agents and without marginal control benefit.

In our model the large shareholder never accumulates shares in equilibrium while this is a possible outcome in the DM-U model. Since the aggregate small investor in the DM-U model is likely to have a rather small coefficient of risk aversion, this difference between the two models mainly comes from the existence in the DM-U model of benefits of control which are increasing in the large stockholder's stake in the firm. We would argue that the absence of such control benefits is an inherent feature in a model in which there is separation between ownership and control. The large shareholder is

[17] The term $f_1(e(a,t), a, t)\partial e(a,t)/\partial a$ is not included in the definition of $b(a,t)$ because the envelope theorem applies to their equation (3). This was made possible because of the absence of a strategic effect in their model.

not, in this context, able to expropriate small shareholders. On the contrary, the manager may be able to "expropriate" all shareholders, whether large or small, and this why the large shareholder exerts some effort to monitor her. We do not see overwhelming reasons why the costs of this monitoring effort should decrease with the large shareholder's stake in the firm. Our paper may be viewed as showing that, when ownership and control are exerted by distinct agents, there is a tendency for large shareholders to divest shares, whether gradually or not. Another question is why, when there exist potential benefits from control, they do not choose to manage the firms directly. A likely answer could be their lack of management ability. This is left for further research.

6 Conclusion

We have shown that a large shareholder divests her shares because, in the absence of share trading, there would exist a wedge between her marginal returns on holding these assets and the atomistic investors' valuation of a share. This wedge arises because the atomistic investors free ride on her monitoring effort which is aimed at reducing the manager's opportunistic behavior (such as choosing projects that are more advantageous to him than to the shareholders). As she divests, the manager increases his effort, but in general this is to the detriment of the firm's profit stream. (We show that $W'(\alpha) > 0$ for $\alpha < \alpha_2^*$). This evolution toward a pure managerial firm, in which the owners do not monitor the manager, can be gradual or immediate, depending on the degree of incongruence of the manager's interest to that of the owners.

Our conclusions differ significantly from the original static model of Burkart, Gromb and Panunzi (1997). While BGP correctly pointed to the conflict of interest between the large shareholder and the manager, and offered an appealing formulation of that conflict, they abstracted from rational expectations of small investors which would arise in a dynamic setting. By fully allowing for rational expectations, we obtain some sharp results.

Even in the case of full commitment, so that asset sales take place only once, the optimal commitment level of stock holding by the large shareholder is different from the level derived by BGP in a static model. In the non-commitment case, we show that complete divestment is the ultimate outcome.

One may wonder whether this tendency for disperse ownership would disappear if the owners can propose incentive contracts to the manager. Burkart, Gromb and Panaunzi (1997) have already answered this question by showing that there nevertheless remains in that case some scope for monitoring and a negative relationship between the manager's effort and the large shareholder's stake in the firm.

Our model can be enriched by extension in several directions. First, one can allow the manager to play a dynamic game. If he realizes that the large shareholder may want to divest, and the speed of divestment depends on the expected net income flow at any level of the state variable α, he may have strong incentives to adopt strategies that condition his effort level on the state variable. Second, the impacts of dynamic incentives on the welfare of the small investors could be investigated in a model where they are not simply indifferent between alternative forms of asset holding. These are parts of our future research agenda.

References

[1] Berle, A.A., Means, G.C., 1932. The Modern Corporation and Private Property. Commerce Clearing House, New York.

[2] Burkart, G., Gromb, D., Panunzi, F., 1997. Large shareholders, monitoring and the value of the firm. Quarterly Journal of Economics 112, 693-728.

[3] Coase, R., 1972.Durability and monopoly. Journal of Law and Economics 15, 143-149.

[4] DeMarzo, P. M., Urošević, B., 2006. Ownership dynamics and asset pricing with a large shareholder. Journal of Political Economy 114, 774-814.

[5] Demsetz, H. and Lehn, K., 1985.The structure of corporate ownership : causes and consequences. Journal of Political Economy 93, 1155-1177.

[6] Dockner, E., Jørgensen, S., Long, N.V., Sorger, G., 2000. Differential Games in Economics and Management Science. Cambridge University Press, Cambridge, U.K.

[7] Driskill, R.,1997. Durable goods monopoly, increasing marginal cost, and depreciation. Economica 64, 137-54.

[8] Edelstein, R.H, Sureda-Gomila, A., Urošević,B., Wonder, N., 2010. Ownership dynamics with multiple insiders: The case of REITs. Real Estate Economics 38: 57-90.

[9] Franks, J., Mayer, C., Rossi, S., 2005. Spending less time with the family: the decline of family ownership in the United Kingdom, in: Morck, R. K. (Ed.), A History of Corporate Governace around the World: Family Business Groups to Professional Managers. NBER Books. University of Chicago Press, Chicago, pp. 581-611.

[10] Gomes, A., 2000. Going public without governance: managerial reputation effects. Journal of Finance 55, 615-645.

[11] Gorga, E., 2009. Changing the paradigm of stock ownership from concentrated towards dispersed ownership? Evidence from Brazil and consequences for emerging countries. Northwestern Journal of International Law and Business 29: 439-458.

[12] Gul, F., H. Sonnenschein and R. Wilson, 1986. Foundations of dynamic monopoly and the Coase conjecture. Journal of Economic Theory 39, 155-90.

[13] Kahn, C., 1987. The durable goods monopolist and consistency with increasing costs. Econometrica 54, 274-294.

[14] Karp, L., 1996. Depreciation erodes the Coase conjecture. European Economic Review 40, 473-490.

[15] LaPorta, R. Lopez, de Silanes, F.L, Shleifer, A., 1999. Corporate ownership around the world. Journal of Finance 54, 471-517

[16] Laussel, D., de Montmarin, M., Long, N. V., 2004. Dynamic duopoly with congestion effects. International Journal of Industrial Organization 22, 655-677.

[17] Long, N. V. 2010, A Survey of Dynamic Games in Economics, World Scientific Publising Co, Singapore.

[18] Mikkelson, W., Partch, M., Shah, K., 1997. Ownership and operating performance of companies that go public. Journal of Financial Economics 44, 281-283

[19] Roe, M.,1996. Strong Managers, Weak Owners : The Political Roots of American Corporate Finance. Princeton University Press, Princeton.

[20] Scott, J., 1990. Corporate control and corporate rule : Britain in an international perspective. British Journal of Sociology 40, 351-373.

[21] Stokey, N., 1981.Rational expectations and durable goods pricing. Bell Journal of Economics 12, 112-128.

[22] Subrahmanyam, A., Titman, S.,1999. The going public decision and the development of financial markets. Journal of Finance 54, 1054-1082.

Type 2	Type 3	Probability
(Π, b)	(Π, b)	$\lambda \in (0, 1)$
$(\Pi, 0)$	$(0, b)$	$1 - \lambda$

Table 1

LEGENDS FOR FIGURES.

Figure 1: The three cases

Figure 2: How long it takes to reach $\alpha = 0$

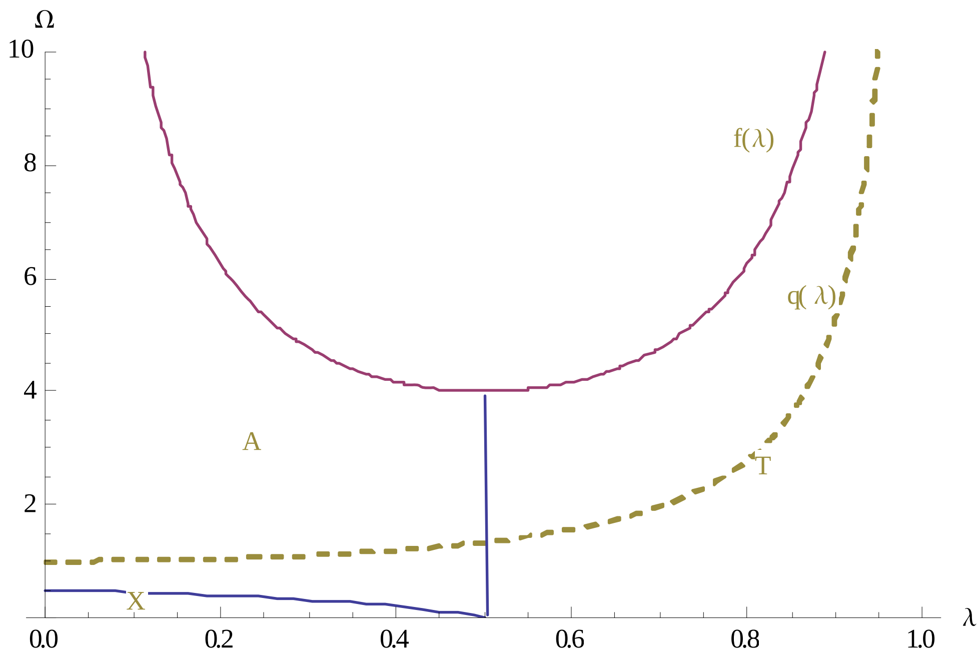

Figure 1: The three cases.

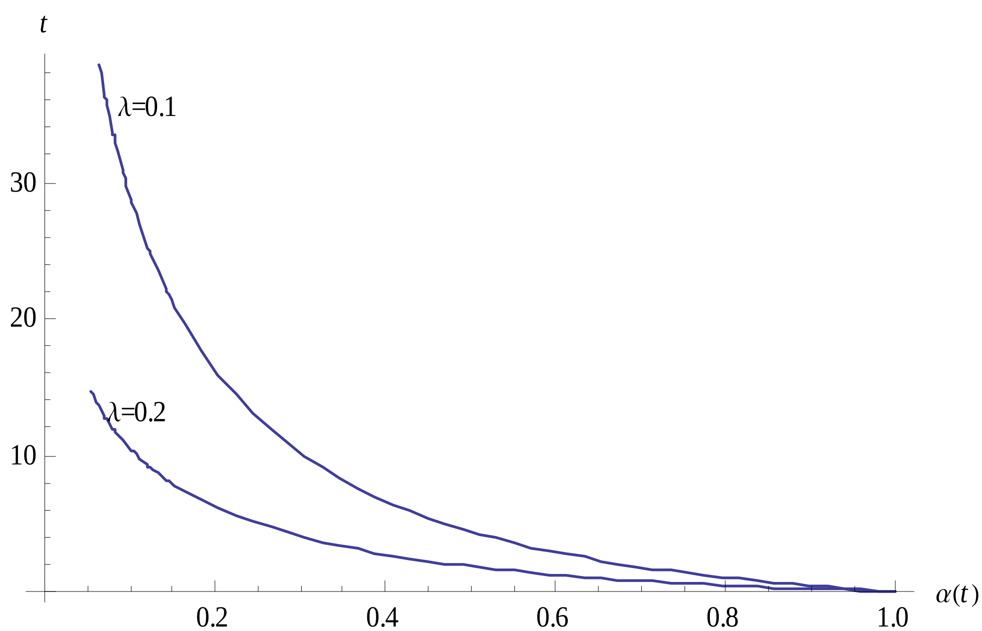

Figure 2: How long it takes to reach $\alpha = 0$.