DISCUSSION PAPER SERIES

No. 3461

OPTIMAL DIVERSIFICATION

Joao Gomes and Dmitry Livdan

FINANCIAL ECONOMICS and INDUSTRIAL ORGANIZATION

ISSN 0265-8003

OPTIMAL DIVERSIFICATION

Joao Gomes, University of Pennsylvania and CEPR
Dmitry Livdan, University of Pennsylvania

Discussion Paper No. 3461
July 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3461

July 2002

ABSTRACT

Optimal Diversification*

In this Paper we show that the main empirical findings about firm diversification and performance are actually consistent with the optimal behaviour of a firm that maximizes shareholder value. In our model, diversification allows a firm to explore better productive opportunities while taking advantage of economies of scale. The dynamic structure of our model allows us to examine several aspects of the relationship between firm diversification and performance in a very general setting.

JEL Classification: D21, G32 and G34
Keywords: corporate strategy, diversification, diversification discount, size and total factor productivity

Joao Gomes
Finance Department
The Wharton School
University of Pennsylvania
Philadelphia
PA 19104-6367
USA
Tel: (1 215) 898 3666
Fax: (1 215) 898 6200
Email: gomesj@wharton.upenn.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=136382

Dmitry Livdan
Finance Department
The Wharton School
University of Pennsylvania
Philadelphia
PA 19104-6367
USA
Email: dlivdan@wharton.upenn.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157907

*We are grateful to comments from Andrew Abel, Michael Brandt, Domenico Cuocco, Francisco Gomes, Gary Gorton, Skander Van den Heuvel, Rich Kihlstrom, Andrew Metrick and Tom Sargent as well as seminar participants at the Wharton School and Penn State.

Submitted 24 June 2002

1 Introduction

Although the degree and type of diversification has evolved over time, diversified firms now account for a large fraction of production and stock market capitalization in most of the developed countries.[1] Yet the research on corporate diversification has been essentially limited to the uncovering of empirical regularities about the effects of diversification on performance. As a consequence, standard microeconomic models largely abstract from the issue altogether and simply assume that firms are homogenous producers specialized in the production of a single good.

In this paper we propose a dynamic model of optimal firm diversification, with roots in the work of Penrose (1959) and Panzar and Willig (1979), where firms seek to maximize shareholder value. In our model, firms diversify for two reasons: (i) to explore new productive opportunities after growth in current operations slows due to decreasing returns; and (ii) to take advantage of synergies across different activities by reducing the fixed costs of production.

The model is well suited to address the main empirical regularities about firm diversification. First, it predicts that diversification is associated with low productivity in on-going activities, as documented by Schoar (2001). Even *before* diversification takes place, a firm that chooses to diversify is less productive than a firm choosing to remain focused. Moreover, since diversification is likely triggered by further reductions in productivity, conglomerates become even less productive *after* they diversify. Thus, even optimal diversification is associated with low productivity, both in the cross-section and in the time series.

Second, decreasing returns on investment mean that *large* firms are more likely to be diversified. This leads to a negative correlation between size and the marginal productivity of capital. Therefore, our model predicts that differences across focused and diversifying firms are, to some extent, related to differences in size, as observed by Santalo (2001).

Finally, despite the fact that diversification maximizes value in our model, we are able to generate the well-documented "diversification discount": a conglomerate is likely to have

[1] This is true both in the US (Montgomery (1994), Villalonga (2001)), Japan (Goto (1981)) and the U.K. (Goudie and Meeks (1982)), for instance.

a lower value of Tobin's Q than comparable stand-alone firms.[2] This is largely due to the selection pattern implied by the optimal diversification decision. A firm that chooses to diversify is not very productive in the first place, while a firm that is productive in its current activities has no interest in dividing resources. This self-selection explanation for the observed discount is consistent with recent work by Chevalier (1999), Villalonga (2001) and Graham, Lemmon and Wolf (2002) that show that the "diversification discount" can actually disappear after controlling for the ex-ante differences across firms.

By emphasizing the role of technologies and costs, our approach to firm diversification also provides a view of conglomerates that contrasts with much of existing literature. Much of the early work on firm diversification has often been interpreted as supporting the view that conglomerates are inefficient, and findings such as the "diversification discount" have generally been explained by appealing to the agency and behavioral problems associated with the existence of conglomerates.[3] However, while focusing on the *effects* of diversification, this literature has generally ignored the *causes* of the diversification decision itself. As a result, this approach raises the more fundamental economic question of why do diversified firms exist at all, if diversification is ex-ante inefficient? By contrast, in our model, although diversified firms are also less efficient ex-post, this result is an endogenous outcome of an optimal diversification strategy and not a direct consequence of an ex-ante inefficiency associated with conglomerates.

A few other authors have also pursued a value maximizing approach to corporate diversification. Maksimovic and Phillips (2002) build a static model of optimal allocation of resources within conglomerates and show that such allocation depends on the relative productivities in different segments. Matsusaka (2001) models diversification as an intermediate, and less productive, stage in a search process over industries that best match the firm's organizational capabilities. When the perfect match is found, a firm specializes. Finally, Bernardo and Chowdhry (2002) justify the existence of diversification discount by assuming that specialized firms have growth options that potentially allow them to diversify

[2]See Wernefelt and Montgomery (1988), Lang and Stulz (1994), Berger and Ofek (1995), Campa and Kedia (1999), Rajan, Servaes, and Zingales (2000), Graham, Lemmon, and Wolf (2001), and Lamont and Polk (2001, 2002) among others.

[3]See Jensen (1986), Amihud and Levy (1981), Jensen and Murphy (1990), Shleifer and Vishny (1989), and Stulz (1990), Denis, Denis and Sarin (1997), and Scharfestein and Stein (2000) among others.

in the future, while conglomerates no longer have these growth options.

While our dynamic environment incorporates a few features from each of these models, our analysis differs in two crucial ways. First, while all of them are designed to address a single specific issue, we are able to provide a unified and consistent explanation for much of the empirical evidence on firm diversification. Our environment is specifically built to deal with the broad array of available evidence by endogenously linking productivity, size, and valuations to diversification strategies. Second, we are able to generate a "diversification discount" endogenously and not by assuming that diversification is ex-ante less valuable. In our model this result stems from endogenous firm selection, an explanation that seems consistent with the recent empirical evidence.

The rest of this paper is organized as follows. Section 2 describes our economic environment and discusses our main assumptions. Section 3 characterizes the optimal strategy of the firm and, in particular, its diversification decision. Section 4 establishes our main empirical implications and Section 5 provides some final remarks.

2 Model

This section describes the optimal behavior of a single firm that is faced with the decision to be diversified or not. To understand its behavior we consider a stylized environment where production of a single good can take place in two separate industries or sectors. Although the model can easily be extended to a larger number of industries, most of its current simplicity would be lost. Moreover, as Lang and Stulz (1994) note, the loss in value associated with diversification is caused by firms going from one to two segments, while the losses due to increasing the number of segments beyond that are insignificant.

2.1 Firm Behavior

Consider the problem of a firm that can choose to operate in two separate sectors, $s = 1, 2$. With a few restrictions, to be detailed below, the firm may engage in production in either segment of the economy at any point in time. Depending on each industry's anticipated profitability, the firm may be active in sector 1, 2 or both. We will say that the firm is diversified if it operates in both sectors simultaneously, and specialized if it is focused on one sector alone.

Production in either sector requires two inputs: capital, k^s, and labor, l^s, and is subject to a sector-specific technology shock, z^s. Labor is hired at the (constant) wage rate $W > 0$. The space of inputs is a subset of the space of (non-negative) real numbers, $K \times L \subseteq R_+^2$. The stochastic process for the shock has a bounded support $Z = [\underline{z}, \overline{z}] \times [\underline{z}, \overline{z}]$, $-\infty < \underline{z} < \overline{z} < \infty$. Moreover define $\Im_z$ and $\Im_k$ as the minimal sigma-fields generated by Z and K, respectively.

Production in each sector, s, is carried out by according to the production function $F : K \times L \rightarrow R_+$

$$y_t^s = z_t^s F(k_t^s, l_t^s). \tag{1}$$

Assumptions 1 and 2 summarize our restrictions regarding the nature of the production function and the stochastic process for productivity.

Assumption 1 *The production function $F(\bullet)$: (i) is continuously differentiable; (ii) is strictly increasing; (iii) is strictly concave; (iv) satisfies the standard Inada conditions; and (v) exhibits decreasing returns to scale in k and l.*

Assumption 2 *The technology levels $z_t = (z_t^1, z_t^2)$ follow a joint Markov transition function $Q(z_{t+1}, z_t) : Z \times \Im_z \rightarrow [0, 1] \times [0, 1]$ that: (i) is stationary, (ii) is monotone and (iii) satisfies the Feller property. Let $G(z)$ denote the invariant distribution of z.*

While most of these assumptions are only imposed for technical reasons, two of them play an important role on the optimal diversification strategy of the firm.

First, decreasing returns to scale and bounded productivity (from the stationarity of $Q(\cdot)$) imply that profit opportunities in each sector are also bounded. Hence, returns to capital accumulation for a focused firm will fall as the firm grows until there is no further incentive for expansion within the current sector. This creates something like a "free cash flow" effect, where available funds are better used by exploring investment opportunities in new segments. This mechanism provides a powerful motivation for firm diversification in our model and is precisely the one explored empirically by Maksimovic and Phillips (2002) and Cocco and Mahrt-Smith (2001).

Second, as long as shocks to the technology levels z_t^1 and z_t^2 are not perfectly correlated, firm diversification will also reduce the variance of cash-flows, providing the firm with another incentive for diversification.

Finally, we assume that production in either sector requires payment of a fixed cost of production, $f \geq 0$. This cost must be paid each period the firm operates in the industry. Diversified firms operate in both sectors but they save a fraction, $0 \leq \lambda/2 \leq 1$, of the costs of operating both technologies. This dilution of the fixed costs due to the presence of synergies and the elimination of corporate redundancies is often cited as one of the major benefits of corporate diversification. These economies of scale provide a final incentive for diversification in our model. We make one final assumption regarding the magnitude of these fixed costs.[4]

Assumption 3 *The fixed costs of production, f, are not too large, i.e. $\exists k \in R_+ : f \leq \underline{z}F(k, l)$.*

Given the above assumptions, current profits for the firm when it specializes in sector s are given by the expression:

$$\pi(s_t, k_t, z_t) = \max_{l_t^s} \left\{ z_t^s F(k_t, l_t^s) - W l_t^s - f \right\}, \tag{2}$$

where W denotes the (constant) wage in terms of final goods and we use $k_t^s = k_t$. If the firm chooses to become a conglomerate, operating in both sectors simultaneously, current profits are:

$$\begin{aligned} \pi(3, k_t, z_t) &= \max_{\left\{l_t^1, l_t^2, \theta_t\right\}} \left\{ z_t^1 F(k_t^1, l_t^1) + z_t^2 F(k_t^2, l_t^2) - W(l_t^1 + l_t^2) - (2 - \lambda) f \right\}, \\ s.t. \qquad 0 &\leq \theta_t \leq 1 \end{aligned} \tag{3}$$

where we define the capital allocations as $k_t^1 = \theta_t k_t$ and $k_t^2 = (1 - \theta_t) k_t$, and the final term in (3) recognizes the cost savings associated with firm diversification and is discussed above.

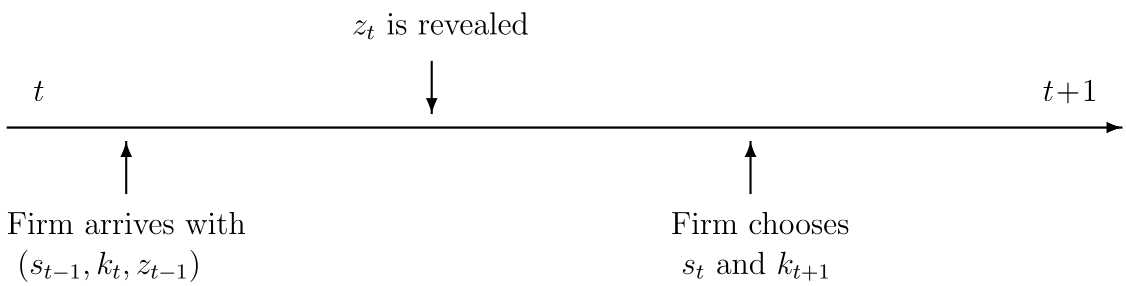
The existence of the profit function, $\pi(s, k, z)$, is guaranteed by Assumption 1. Moreover, it also follows immediately that $\pi(s, k, z)$ is: (i) bounded, continuously differentiable, strictly

[4]Fixed costs also guarantee a minimum scale of production, thus forcing a firm to stay focused, unless outside opportunities appear attractive enough. As we show below, given decreasing returns to scale and without fixed costs a firm will always be diversified.

increasing, and strictly concave in $k \in K$, (ii) bounded, continuous and strictly increasing in $z \in Z$.

The dynamic nature of the model is captured by two endogenous transitions, regarding capital accumulation and the optimal sectoral choices of the firm. The timing of the decisions is illustrated in Figure 1 and is as follows. At the beginning of every period t the firm chooses: (i) whether to operate one or both production units; (ii) the optimal allocation of inputs across its different activities; and (iii) how much to invest for the future.

Figure 1: Timing of Events



Sectoral choices are constrained so that firms cannot simply "jump" across sectors, a pattern not observed in the data. To accomplish this we assume that if the firm was previously focused in sector s, it can only choose between remaining focused in the same sector, or diversify and operate both production units. Diversified firms, on the other hand, can choose to remain diversified, or specialize in either sector 1 or 2. We will also refrain from modelling entry and exit decisions. Entry/exit will only complicate the model without providing any new significant insight, beyond what our current setup can do, on the problem at hand.

Finally, the evolution of the capital stock of the firm is described by the law-of-motion

$$k_{t+1} = (1 - \delta)k_t + i_t, \tag{4}$$

where i_t denotes gross investment spending, and δ is the rate depreciation of capital.

2.2 Discussion

Our environment incorporates three of the most popular reasons for diversification to be *optimal* strategy: the existence of synergies, the availability of "free" cash flows and pooling of risk. Synergies are created by elimination of redundancies across business lines, such as overhead. In our model, we capture this feature by introducing the savings parameter λ. "Free" cash flows result from a firm's excess capacity in valuable resources, that can be transferred to other activities. As we have seen, decreasing returns to scale provide us with a convenient way of formalizing this concept. As the firm grows in size, marginal productivities fall and additional resources can be better allocated to new activities. This feature of our environment is also consistent with the empirical observation that large firms are much more likely to become diversified than smaller ones. Finally, as long as the two shocks z_t^1 and z_t^2 are not identical firm diversification will also reduces cash flow risk.

Clearly, in the absence of trading frictions, risk pooling is not valued by investors, since they can hedge this risk on their own. Synergies and decreasing returns, however, generate value to shareholders. In both cases, production is more efficient and resources are saved when operations are combined in a conglomerate.

Finally, as in Bernardo and Chowdhry (2002), our model also incorporates an option to diversify in the valuation of a focused firm. In their work this is the key assumption to generate the diversification discount. Here, however, diversification is not irreversible and the diversified firm has the option to refocus in the future. Interestingly, this option is actually more valuable since we assume that its segment choices are unconstrained.

To summarize we have introduced a fairly comprehensive model of dynamic firm behavior, which incorporates several plausible benefits to corporate diversification while abstracting from any of its potential drawbacks, such as those induced by agency problems. Nevertheless, as we shall see, it is quite possible to find evidence of a diversification discount in this environment.[5]

[5]explored in the literature such as greater debt capacity (Lewellen (1971)), monopoly power (Tirole (1995)), and the existence of internal capital markets (Alchian (1969) and Williamson (1975)).

2.3 Optimality

As Figure 1 shows, the firm's decisions at any period t are conditional on the current values of the exogenous vector of productivity shocks z_t, and on its beginning of period size, k_t. Moreover, its sectoral choices depend on its previous activities, s_{t-1}. To save on notation, we will use $s = 3$ to denote a diversified firm and define the set $S = \{1, 2, 3\}$. Accordingly, let the vector $(s, k, z) \in S \times K \times Z$ denote the current state for a firm that was active in sector s during last period, has k units of physical capital installed and faces a vector of technology shocks z. As we have seen, this completely characterizes the information available to the firm at the beginning on the period, before any decisions are made.

Given the recursive structure of the environment it is convenient to formulate the firm's problem using dynamic programming. Assuming that the firm maximizes the present discounted value of future cash flows, its optimal market value, $v(s, k, z)$, solves the following Bellman equation:

$$\begin{aligned} v(s,k,z) &= \max_{\{k',s'\}} \left\{ \pi(s',k,z) + (1-\delta)k - k' + \beta \int v(s',k',z')Q(dz',z) \right\} \qquad (5) \\ s' &\in \begin{cases} \{s,3\}, & s = 1,2 \\ S, & s = 3 \end{cases} \end{aligned}$$

where $0 < \beta < 1$ is the discount factor for the owner of the firm.[6]

There are two components to the right hand side of the (5). The first term describes the current period payoffs, equal to profits, defined by (2) and (3), minus investment spending, described in (4). Note that current period profits depend on the current sectoral choice of the firm, $s' = s_{t+1}$. The last term captures the (discounted) expected continuation value to the firm, which depends on the current decisions about capital accumulation, k', sectoral choice, s', as well as the future productivity shocks z'. Finally, as discussed above, the current sectoral choice of a previously focused firm, is constrained by a "no-switching" rule: no firm can jump directly across sectors.

Proposition 1 establishes the existence of a unique function $v(s, k, z)$, that satisfies (5), while Proposition 2 lists some of its basic properties.

[6] To save notation we will use the convention that k', z', and s' denote the value of the state variables at the beggining of the next period.

Proposition 1 *There exists a unique function $v(s, k, z)$ that solves the dynamic program (5).*

Proof. See Appendix A. ■

Proposition 2 *The value function $v(s, k, z)$, defined in (5) is: (i) continuous; and (ii) increasing in both k and z.*

Proof. See Appendix A. ■

Note that the value function is always increasing in the vector of shocks $z = (z^1, z^2)$. In other words, the value of the firm increases in each shock, regardless of whether the firm was operating in that sector or not.

3 Optimal Corporate Strategy

In this section we use our model to explicitly characterize the optimal corporate strategies of the firm. In our model, these strategies are summarized by a set of functions $k' = \mathbf{k}(s, k, z)$, and $s' = \mathbf{s}(s, k, z)$, that characterize both the optimal size of the firm and the segments in which it operates. We start by formally defining these optimal decisions and then proceed to examine their basic properties.

3.1 Firm Decisions

Since S is a set of discrete numbers, the optimal industrial decision, $s' = \mathbf{s}(s, k, z)$, can be computed as follows. First, define the function

$$p(s', k, z) \equiv \max_{k'} \left\{ \pi(s', k, z) + (1 - \delta)k - k' + \beta \int v(s', k', z') Q(dz', z) \right\}. \tag{6}$$

as the value of the firm, conditional on having taken the sectoral decision s' in the current period.

The exact condition for optimal diversification depends on the number of activities of the firm at the beginning of the period, s, since focused firms are not allowed to simply switch industries. Hence for a firm previously specialized in sector $s \in \{1, 2\}$, diversification is optimal if

$$p(3, k, z) \geq p(s', k, z) \mid_{s'=s} = p(s, k, z) \tag{7}$$

while a previously diversified firm ($s = 3$) will remain diversified only if

$$p(3,k,z) \geq \max\{p(1,k,z), p(2,k,z)\} \tag{8}$$

Hence, the optimal choice of segments can be summarized as:

$$\mathbf{s}(s,k,z) = \begin{cases} 1, \text{ if } & \begin{cases} p(1,k,z) > p(3,k,z) \text{ and } s = 1 \\ p(1,k,z) > \max\{p(2,k,z), p(3,k,z)\} \text{ and } s = 3 \end{cases} \\ 2, \text{ if } & \begin{cases} p(2,k,z) > p(3,k,z) \text{ and } s = 2 \\ p(2,k,z) > \max\{p(1,k,z), p(3,k,z)\} \text{ and } s = 3 \end{cases} \\ 3, \text{ else} & \end{cases}, \tag{9}$$

The optimal capital accumulation decision can be characterized in the same way. First, define the optimal choice of capital conditional on the contemporaneous choice of segment s', as[7]

$$k^*(s',z) \equiv \arg\max_{k'} \left\{ \beta \int v(s',k',z')Q(dz',z) - k' \right\}. \tag{10}$$

The optimal capital accumulation is then defined as:

$$\mathbf{k}(s,k,z) \equiv k^*(\mathbf{s}(s,k,z),z) \tag{11}$$

Finally, it is worth noting that the value of the firm satisfies

$$v(s,k,z) = p(\mathbf{s}(s,k,z),k,z) = \begin{cases} \max\{p(s,k,z), p(3,k,z)\}, & s \in \{1,2\} \\ \max\{p(1,k,z), p(2,k,z), p(3,k,z)\}, & s = 3 \end{cases}. \tag{12}$$

which implies that

$$\begin{aligned} v(3,k,z) &= \max\{\max\{p(1,k,z), p(3,k,z)\}, \max\{p(2,k,z), p(3,k,z)\}\} \\ &= \max\{v(1,k,z), v(2,k,z)\}. \end{aligned} \tag{13}$$

It follows that the value of a previously diversified firm is always at least as large as that of a previously focused firm.

3.2 Optimal Diversification

Definitions

With these definitions at hand we can now examine the properties of the optimal diversification strategies (7) and (8) in detail. Combining equations (6) and (7) we can

[7] Note that, conditional on the number of segments that the firm chooses to operate this period, s', optimal investment is independent of current size, k. Intuitively, without adjustment costs, capital always adjusts immediately to its optimal level.

decompose the diversification decision for a previously specialized firm, as follows:

$$p(3,k,z) \geq p(s,k,z) \Leftrightarrow \Pi(s,k,z) + \Psi(s,z) \geq (1-\lambda)f, \quad s \in \{1,2\} \tag{14}$$

where

$$\Pi(s,k,z) \equiv \pi(3,k,z) - \pi(s,k,z) + (1-\lambda)f, \tag{15}$$

and

$$\begin{aligned} \Psi(s,z) &= \Psi(s',z) \mid_{s'=s} \equiv \max_{k'} \left\{ \beta \int v(3,k',z')Q(dz',z) - k' \right\} - \\ &-\max_{k'} \left\{ \beta \int v(s,k',z')Q(dz',z) - k' \right\}. \end{aligned} \tag{16}$$

Equation (14) decomposes the optimal diversification decision into two parts: a "profit" component, $\Pi(s,k,z)$, that results from comparing current period payoffs of the alternative strategies, and an "option" component, $\Psi(s,z)$, associated with the continuation payoffs.

Similarly, the optimal diversification decision for the previously diversified firm can be written as

$$p(3,k,z) \geq \max\{p(1,k,z), p(2,k,z)\} \Leftrightarrow \min_{s\in\{1,2\}} \{\Pi(s,k,z) + \Psi(s,z)\} \geq (1-\lambda)f. \tag{17}$$

Lemma 3 summarizes the basic properties of the profit effect. Since the diversified firm can always replicate a focused firm by allocating all capital to the relevant industry, its profits (absent the fixed costs) can never be smaller. In addition, by operating in two separate industries, a diversified firm will also have a higher return on capital. Finally, the profits for a focused firm are more vulnerable to shocks to its current line of business.

Lemma 3 *Let $\Pi(s,k,z)$ be defined by (15). Then $\Pi(s,k,z)$ is (i) non-negative; (ii) weakly increasing in k; and (iii) decreasing in z^s and increasing in $z^{\widetilde{s}}$, $s \neq \widetilde{s}$.*

Proof. See Appendix A ■

Lemma 4 contains the main properties of the "option" effect. Intuitively, given the unrestricted options of diversified firms, this effect cannot be negative. In addition, and just as with profits, the continuation value of a specialized firm is closely tied to its on-going activities and thus is more sensitive to their fluctuations.

Lemma 4 *Let $\Psi(s, z)$ be defined by (16). Then $\Psi(s, z)$ is: (i) non-negative; and (ii) decreasing in z^s and increasing in $z^{\tilde{s}}$, $s \neq \tilde{s}$.*

Proof. See Appendix A ■

Diversification Threshold

Having established these basic results we can now examine the main properties of the optimal diversification decision of the firm. We first show that this decision can be summarized by something like an "indifference curve", or, alternatively, a "diversification threshold", in the space of state variables.

Proposition 5 *The optimal diversification decision can be characterized by the unique threshold value:*

$$\hat{k}(s, z) = \arg\min_{k} \{\mathbf{s}(s, k, z) = 3\}, \quad \forall (s, z) \in S \times Z \tag{18}$$

Proof. Consider the case of a previously specialized firm. The optimal diversification threshold satisfies

$$\Pi(s, \hat{k}(s, z), z) + \Psi(s, z) = (1 - \lambda) f, \quad \forall (s, z) \in S \times Z \tag{19}$$

Lemmas 3 and 4 imply that the left hand side is both non-negative and strictly increasing in k. Hence, if diversification is optimal for $k = \hat{k}(s, z)$, it must also be optimal for $k > \hat{k}(s, z)$. If the left hand side exceeds $(1 - \lambda) f$ then diversification is always optimal and $\hat{k}(s, z) = 0$.

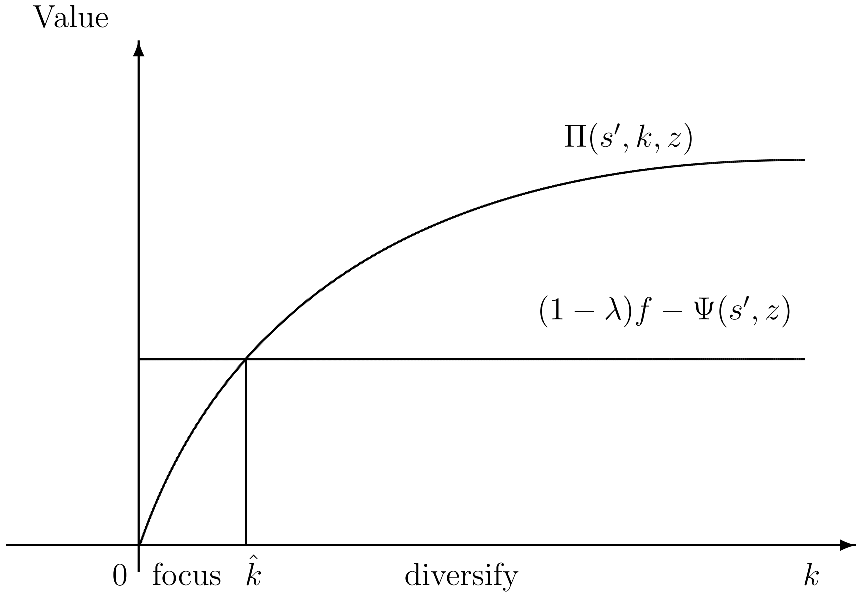
Now consider a previously diversified firm. Here the threshold is determined by

$$\min_{s \in \{1,2\}} \left\{ \Pi(s, \hat{k}(s, z), z) + \Psi(s, z) \right\} = (1 - \lambda) f, \quad \forall z \in Z \tag{20}$$

Again the left hand side is both non-negative and strictly increasing in k, since $\Pi(s, k, z) + \Psi(s, z)$ has these properties as well, and the result follows as above. ■

Figure 2 illustrates the determination of this optimal diversification threshold, $\hat{k}(s, z)$, for a firm with a productivity vector z, which was previously engaged in segment s.

Figure 2: Computing the Diversification Threshold



The intuition for the result in Proposition 5 is fairly simple but quite important. Diversification results from comparing current profits against expected continuations. As we have already seen, fixed costs aside, current profits must be larger for a diversified firm since it can allocate capital to take better advantage of production opportunities in two sectors. The same is true of the continuation value, $\Psi(s, z)$. Since a diversified firm has the option to retrench into any sector, $\Psi(s, z)$ can never be negative.

This implies that there is only one reason for the firm to remain focused: the higher fixed costs associated with operating two separate technologies. As long as $(1 - \lambda)f$ is high enough small firms will find it unprofitable to split their capital, since they cannot cover these additional production costs. This result seems to agree with the empirical findings that large firms are more likely to be diversified.

Corollaries 6 and 7 establish two properties of the optimal industrial strategy $\mathbf{s}(s, k, z)$. Corollary 6 shows why the role of fixed costs is crucial in our analysis. Without them, profits are always positive in both sectors and the firm would have no incentive to focus, given the assumption of decreasing returns to scale. Corollary 7 shows that if synergies are sufficiently large there is never an incentive for the firm to be focused.

Corollary 6 *In the absence of fixed costs ($f = 0$), diversification is the optimal corporate strategy.*

Proof. In the absence of fixed costs inequality (14) is always satisfied. ■

Corollary 7 *Suppose $f > 0$. Diversification is the optimal corporate strategy if $\lambda \geq 1$, i.e. synergies are sufficiently large.*

Proof. Inequality (14) is always satisfied if $\lambda \geq 1$. ■

Corollary 8, uses the complete symmetry between sectors in our economy, to establish a useful property of the diversification threshold, $\hat{k}(s, z)$.

Corollary 8 *Assume $z^1 = z^2$. Then $\hat{k}(1, z) = \hat{k}(2, z)$.*

Proof. This follows immediately from the symmetry of the problem. ■

Finally, we also show that the optimal threshold for a previously diversified firm, $\hat{k}(3, z)$, is completely characterized by the threshold rules for previously focused firms $\hat{k}(1, z)$ and $\hat{k}(2, z)$.

Proposition 9 *The optimal threshold for a previously diversified firm, $\hat{k}(3, z)$ satisfies*

$$\hat{k}(3, z) = \max_{s \in \{1,2\}} \hat{k}(s, z), \qquad \forall z \in Z$$

Proof. Consider the case where

$$\Pi(1, \hat{k}(3, z), z) + \Psi(1, z) < \Pi(2, \hat{k}(3, z), z) + \Psi(2, z)$$

then (20) implies that

$$\Pi(2, \hat{k}(1, z), z) + \Psi(2, z) > \Pi(1, \hat{k}(1, z), z) + \Psi(1, z) = (1 - \lambda)f$$

and $\hat{k}(3, z) = \hat{k}(1, z)$. But from (18) we know that

$$\Pi(2, \hat{k}(2, z), z) + \Psi(2, z) = (1 - \lambda)f$$

since $\Pi(2, k, z)$ is increasing in k, it follows that

$$\hat{k}(3, z) = \hat{k}(1, z) > \hat{k}(2, z)$$

The other case follows immediately. ■

4 Empirical Implications

Despite its simplicity and generality, our model provides a useful framework to analyze a number of empirical issues regarding the impact of diversification on firm performance. In this section we address two such issues: the effects of diversification on productivity, and, the "diversification discount".

4.1 Diversification and Productivity

The implications of diversification for productivity have been empirically examined by Lichtenberg (1992), and recently by Maksimovic and Phillips (2001, 2002) and Schoar (2001). All of these authors use the Longitudinal Research Database provided by the Bureau of Census. While Lichtenberg's (1992) results are fairly ambiguous, both Schoar and Maksimovic and Phillips (2001, 2002) find that acquired assets experience increase in productivity after the ownership change. While we cannot explicitly deal with this evidence in the context of our model note that this is clearly supportive of our optimal view of diversification. More importantly however, Schoar (2001) also finds strong evidence that diversifying firms experience a noticeable drop in productivity in incumbent plants, i.e. those operated prior to the expansion. This effect actually dominates the gains in the new plants and results in a net drop in firm productivity following diversification. This finding is usually presented both as evidence, and as an explanation, for the traditional view that diversification destroys value. In this section we show that this result is actually quite consistent with our view of diversification as a value maximizing strategy.

To accomplish this, we first establish a few additional properties of the diversification threshold $\hat{k}(s,z)$. These are summarized in Proposition 10 below.

Proposition 10 *Let* $z = (z^s,\, z^{\widetilde{s}})$ *and* $\{s, \widetilde{s}\} = 1, 2$. *The diversification threshold,* $\hat{k}(s,z)$, *is: (i) increasing in* z^s *and, (ii) decreasing in* $z^{\widetilde{s}}$.

Proof. Let $z = (z^s, z^{\widetilde{s}})$ and $\widehat{z} = (z^s + \Delta z^s, z^{\widetilde{s}})$, with $\Delta z^s > 0$. It follows from equation (19) that

$$\Pi(s, \hat{k}(s, \widehat{z}), \widehat{z}) + \Psi(s, \widehat{z}) = (1 - \lambda) f.$$

Lemmas 3 and 4 imply that both $\Pi(\cdot)$ and $\Psi(\cdot)$ are decreasing in z^s. Since $\Pi(\cdot)$ is increasing in k (also Lemma 3), it follows that $\hat{k}(s, \widehat{z}) > \hat{k}(s, z)$.

Analogously, let $\widehat{z} = (z^s, z^{\widetilde{s}} + \Delta z^{\widetilde{s}})$, with $\Delta z^{\widetilde{s}} > 0$. Since both $\Pi(\cdot)$ and $\Psi(\cdot)$ are increasing in $z^{\widetilde{s}}$ (Lemmas 3 and 4), it follows that $\hat{k}(s, \widehat{z}) < \hat{k}(s, z)$. ■

Corollary 11 *Let $z = (z^s, z^{\widetilde{s}})$. The diversification threshold $\hat{k}(3, z)$ is: (i) increasing in z^s and decreasing in $z^{\widetilde{s}}$, for $z^s > z^{\widetilde{s}}$; and, (ii) decreasing in z^s and increasing in $z^{\widetilde{s}}$, for $z^s < z^{\widetilde{s}}$.*

Proof. Suppose $z^s > z^{\widetilde{s}}$. It follows from Corollary 8 and Proposition 10 that

$$\hat{k}(3, z) = \hat{k}(s, z) > \hat{k}(\widetilde{s}, z). \tag{21}$$

It follows from Proposition 5 that $\hat{k}(3, z)$ is increasing in z^s and decreasing in $z^{\widetilde{s}}$. Now suppose $z^s < z^{\widetilde{s}}$, then

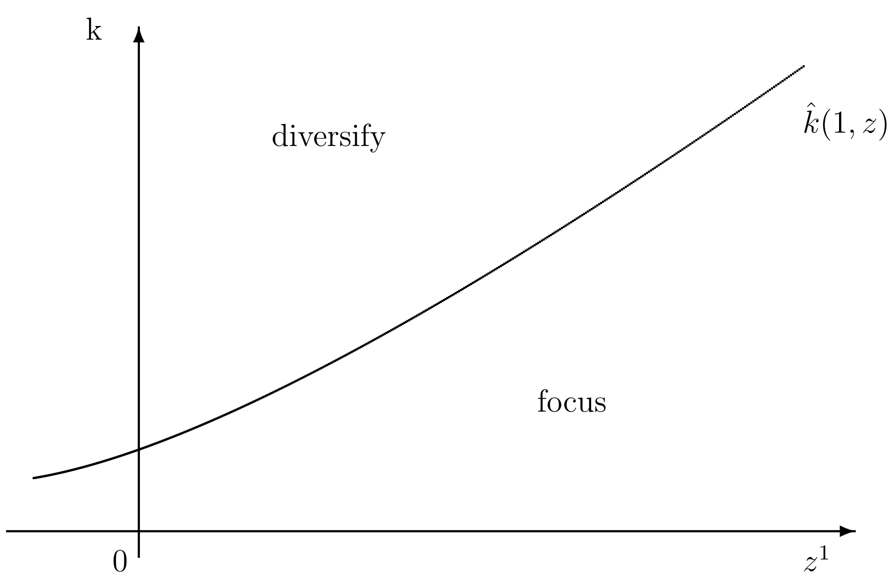
$$\hat{k}(3, z) = \hat{k}(\widetilde{s}, z) > \hat{k}(s, z).$$

and it follows that $\hat{k}(3, z)$ is decreasing in z^s and increasing in $z^{\widetilde{s}}$. ■

Remark 12 *For a fixed $z^{\widetilde{s}}$, the diversification threshold $\hat{k}(3, z)$ reaches a minimum at $z^s = z^{\widetilde{s}}$.*

The results of Proposition 10 are illustrated in Figure 3. It plots the diversification threshold for a firm previously specialized in sector 1, $\hat{k}(1, z)$.

Figure 3: The Diversification Threshold



Remember that a firm entering the current period as focused in sector 1 has only two choices: it can either remain in sector 1, or diversify and operate in both sectors simultaneously. As we have seen above, the firm's decision will depend on size, k, and the value of productivity in both sectors, z^1 and z^2. The positive slope of $\hat{k}(1, z)$, implies that, given their size, firms are more likely to remain focused when productivity is high in the incumbent sector, z^1, while diversification becomes optimal when this productivity becomes too low.

This result has two very powerful implications for the empirical literature on the relation between diversification and productivity. First, from a dynamic, or time series, standpoint, diversification in our model is associated with lower productivity in incumbent operations, as documented by Schoar (2001). Since, controlling for size, diversification occurs only when z^1 falls below the threshold (given by the inverse of $\hat{k}(1, z)$), ex-post total factor productivity must be lower for diversifying firms. Second, from a static, or cross-sectional, viewpoint, since firms closer to the threshold are more likely to diversify, it follows that, controlling for size, their productivity must be lower than that of those focused firms less likely to diversify. Hence, ex-ante productivity for diversifying firms is also below that of firms that choose to remain focused firms.

Productivity and Size

Although we have focused on the role of productivity shocks holding size fixed, it is clear that the optimal diversification decision also depends on the size of on-going operations. The positive slope of $\hat{k}(1, z)$, implies that, for a given productivity vector, z, "small" firms are much less likely to diversify than "large" firms. In this sense our model can produce a "size" effect, documented by, among others, Santalo (2001). Given our assumption of decreasing marginal returns, size is negatively related to the marginal productivity of capital. Hence, the "size" effect will reinforce our findings that both ex-ante and ex-post productivity are lower for diversifying firms. Note, however, that these results also imply that the "size" effect does not account for all of the effects of diversification on productivity.

4.2 The Diversification Discount

A significant number of studies find that conglomerates have a lower value for Tobin's Q and trade at a discount relative to a replicating portfolio of stand-alone firms.[8] While this puzzling result is usually attributed to the fact that conglomerates are inefficient and value destroying, we show in this section that it can be rationalized in our model of optimal diversification as well.

Conditional Discount

Let θ^* be the optimal share of capital allocated to sector 1 by the conglomerate. We can then construct a replicating portfolio by splitting this firm into two stand-alone units of size $\theta^* k$ and $(1-\theta^*)\,k$. The "diversification discount" can then be stated as:

$$D(z) = p\,(1, \theta^* k, z) + p\,(2, (1-\theta^*)\,k, z) - p\,(3, k, z) > 0 \tag{22}$$

In other words, the market value of the replicating portfolio exceeds that of the conglomerate, after conditioning for firm characteristics, in this case size and productivity. We can use the definition of Tobin's average q, for a firm that follow strategy $s' = \mathbf{s}\,(s, k, z)$,

$$q\,(s', k, z) \equiv \frac{p\,(s', k, z)}{k}, \quad \forall s' \in S, \forall (k, z) \in K \times Z \tag{23}$$

and rearrange (22) to obtain an alternative formulation of the diversification discount:

$$D(z) = \left(\theta^* q\,(1, \theta^* k, z) + (1-\theta^*)\,q\,(2, (1-\theta^*)\,k, z)\right) - q\,(3, k, z) > 0$$

Hence the diversification discount can also be stated as the difference between the weighted average of q's for the stand alone divisions and the value of Tobin's q for the conglomerate (See Lang and Stulz (1994)).

While this puzzling result is usually attributed to the fact that conglomerates are inefficient and value destroying, we show below that it can be rationalized in our model of optimal diversification as well. To prove this result we need to make use of the following lemma.

[8]See Wernefelt and Montgomery (1988) and Lang and Stulz (1994) for example.

Lemma 13 *Define the function:*

$$C(z) = \max_{k'}\left\{\int v(1,k',z')Q(dz',z) - k'\right\} + \max_{k'}\left\{\int v(2,k',z')Q(dz',z) - k'\right\} \quad (24)$$
$$-\max_{k'}\left\{\int v(3,k',z')Q(dz',z) - k'\right\}, \qquad \forall s' \in S, \forall (k,z) \in K \times Z$$

Then $C(z) \geq 0$, $\forall z \in Z$.

Proof. See Appendix A. ■

Proposition 14 shows that a diversification discount (conditional on both size and productivity) exists if the synergies from diversification are not too large.

Proposition 14 *Diversification leads to a discount if*

$$\lambda f \leq C(z)$$

Proof. Note that (3) implies that

$$\pi(3,k,z) = \pi(1,\theta^* k,z) + \pi(2,(1-\theta^*)k,z) + \lambda f$$

Using in valuation equation (6), it follows that the discount (22) can be written as:

$$D(z) = p(1,\theta^* k,z) + p(2,(1-\theta^*)k,z) - p(3,k,z) = C(z) - \lambda f \geq 0$$

■

The intuition for this striking result is quite simple. The basic benefit of diversification is the creation of synergies that lower the fixed costs of production by λf. Diversification has an important cost, however, when compared with a portfolio of specialized firms: while the continuation value for a diversified firm is at least as high as that of the largest specialized firm (equation (13)) it is simply smaller than the sum of the individual continuation values for both stand-alone firms.[9]

Empirically, the fact that the discount depends on the extent of the synergies appears quite plausible. Synergies are likely high when firms diversify into related industries, and there is an ample empirical support for the view that related diversification leads to better performance and a lower (or non-existent) diversification discount (for example Chevalier (2000)).

[9] Alternatively, the diversified firm has less "flexibility" in its future decisions (Lemma 13), than a comparable portfolio of firms.

The Unconditional Discount

With a few exceptions, however, most measures of the diversification discount focus on the following cross-sectional regression:

$$q_i = \alpha_0 - \alpha_1 d_i + \alpha_2 \log(k_i), \qquad (25)$$

where q_i is the value of Tobin's q, and d_i is a diversification dummy that takes the value of 1 if firm i is diversified and 0 otherwise. The coefficient $\alpha_1 > 0$ gives then an estimate of the discount associated with diversification. Formally, this is equivalent to simply looking at the sample averages of Tobin's q across focused and diversified firms.

As Villalonga (2001) argues, this is only an unconditional measure of the diversification discount, that, therefore, it fails to control effectively for differences in the ex-ante propensities to diversify across firms. Since we explicitly account for the diversification decision of the firm, our model seems naturally suited to address these selection issues.

In our context, as we have already seen, a firm is more likely to diversify when productivity in on-going activities is quite low. Since the value function $v(.)$ is increasing in z (Proposition 2), it follows that $p(\cdot)$, and hence q, is also increasing in productivity. Hence, diversifying firms are also likely have relatively lower (unconditional) valuations, purely due to this selection bias.

Unfortunately, it is not possible to analytically establish the conditions guaranteeing that $\alpha_1 > 0$, and we will need to use a simple numerical example to illustrate this result. Assume that technology is Cobb-Douglas, so that the profit function (2) has the following representation:

$$\pi(s,k,z) = Ae^{z^s}k^\gamma - f, \quad 0 < \gamma < 1, s = 1, 2.$$

where the stochastic process for the productivity shocks is described by

$$\begin{bmatrix} \log z^1_{t+1} \\ \log z^2_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \log z^1_t \\ \log z^2_t \end{bmatrix} + \begin{bmatrix} \varepsilon^1_{t+1} \\ \varepsilon^2_{t+1} \end{bmatrix},$$

and ε^1_t and ε^2_t are jointly normal random variables

$$\begin{bmatrix} \varepsilon^1_t \\ \varepsilon^2_t \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right].$$

Table 1 summarizes our choice of the key parameter values.[10]

TABLE 1

Parameter Values				
γ	f	σ	ρ	λ
0.9	0.1	0.025	0.95	0.5

With the help of these parameter choices, it is straightforward to (numerically) compute the value function $v(s,k,z)$, optimal corporate strategies, $s' = \mathbf{s}(s,k,z)$, and, the implied values of Tobin's q.

With these values at hand we can compute the implied value of the discount α_1 in our model. Table 2 shows that we can generate sizable discounts for several alternative combinations of the most important parameters.

TABLE 2

Diversification	**Discount**
Benchmark	0.227
$\lambda = 0.4$	0.183
$f = 0.125$	0.117

While the exact size of the discount naturally varies with the exact experiment, it is easy to see that it is a relatively natural outcome of our model.

To summarize we find that our model can rationalize the documented, unconditional, diversification discount. To a large extent this is a consequence of a the selection biases introduced by failing to control for systematic productivity differences across firms. In addition, it is also possible to obtain a diversification discount, even if one conditions on all firm characteristics. This happens however only if the synergies associated with the diversification are not too large.

5 Conclusions

In this paper we show that a general dynamic model of optimal behavior of a firm that maximizes shareholder value is actually consistent with the main empirical findings about

[10]The remaining parametesr, β and δ, are not essential to our results. Nevertheless we use fairly standard values with $\beta = 1/1.065$ and $\delta = 0.1$. Finally we set A so that the average size of the firm equals 1.

firm diversification and performance. Here, diversification is a natural result of firm growth and it stems from dynamic firm strategies that maximize value. Diversification allows a firm to explore new productive opportunities, while taking advantage of economies of scale and reducing the volatility of its cash flows.

The dynamic structure of our model allows us to examine several aspects of the relationship between firm diversification and performance in a very general setting. In particular, we need not place any significant restrictions on the nature of functional forms or parameter values in our model, beyond those already discussed. The very forces leading to optimal diversification, are sufficient to generate a wealth of realistic features, regarding firm diversification, size, productivity and valuations.

We obtain several important results. First, we can show that firms currently expanding are not only less productive than other (non-expanding) focused firms, but they also experience productivity losses after the expansion, as documented by Schoar (2001). Second, as Santalo (2001), we find that size differences can account for part of the differences both in productivity and valuation across focused and diversifying firms. However, we also show that this size "effect", can not account for all of these differences. Finally, and perhaps more surprisingly, we show that despite all the obvious advantages to firm diversification and the fact that firm diversification does not destroy value in our model, it is still possible to obtain a diversification discount.

References

[1] Alchian, A.A., 1969, Corporate Management and Property Rights, in H. Manne, ed.: Economic Policy and the Regulation of Corporate Securities. American Enterprise Institute, Washington, D.C., 337-360.

[2] Amihud, Yakov, and Baruch Lev, 1981, Risk Reduction as a Managerial Motive for Conglomerate Mergers, *Bell Journal of Economics* 12, 605-617.

[3] Berger, Philip and Ofek, Eli, 1995, Diversification's Effect on Firm Value, *Journal of Financial Economics* 37, 39-65.

[4] Bernardo, Antonio and Chowdhry, Bhagwan 2002, Resources, Real Options and Corporate Strategy, *Journal of Financial Economics*, forthcoming.

[5] Campa, Jose and Kedia, Simi, 1999, "Explaining the Diversification Discount", unpublished manuscript, Harvard University.

[6] Chevalier, Judith 1999, Why Do Firms Undertake Diversifying Mergers? An Examination of the Investment Policies of Merging Firms, unpublished manuscript, University of Chicago.

[7] Cocco, Joao and Jan Mahrt-Smith, 2001, "Return Sensitivity to Industry Shocks: Evidence on the (In)-Efficient Use of Internal capital Markets, unpublished manuscript, London Business School.

[8] Dennis, Dabid, Dennis, Dianne and Sarin, Atulya, Agency Problems, Equity Ownership and Corporate Diversification, *Journal of Finance*, 52(1), 135-160.

[9] Graham, John, Lemmon, Michael and Wolf, Jack, 2001, Does Corporate Diversification Destroys Value?, unpublished manuscript, Duke University.

[10] Jensen, Michael, 1986, Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, *American Economic Review* 76(2), 323-329.

[11] Jensen, Michael, and Kevin Murphy, 1990, Performance Pay and Top Management Incentives, *Journal of Political Economy* 98, 225-264

[12] Lamont, Owen and Polk, Chistopher 2001, The Diversification Discount: Cash Flows Versus Returns, *Journal of Finance*, 56(5), 1693-1721.

[13] Lamont, Owen and Polk, Chistopher 2002, Does Diversification Destroys Value? Evidence from Industry Shocks, *Journal of Financial Economics* vol. 63(1)

[14] Lang, Larry and Rene Stulz, 1994, Tobin's Q,Corporate Diversification, and Firm Performance, *Journal of Political Economy* 102(6) 1248-1280.

[15] Lewellen, Wilbur, 1971, A Pure Financial Rationale for the Conglomerate Merger, *Journal of Finance* 26, 527-537.

[16] Lichtenberg, Frank, 1992, Industrial De-Diversification and Its Consequences for Productivity, *Journal of Economic Behavior and Organization* 18, 427-438.

[17] Maksimovic, Vojislav and Phillips, Gordon, 2001, The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains, *Journal of Finance,* 61, 2019-2065.

[18] Maksimovic, Vojislav and Phillips, Gordon, 2002, Do Conglomerate Firms Allocate Resources Inefficiently Across Industries, forthcoming, *Journal of Finance* .

[19] Matsusaka, John, 2001, Corporate Diversification, Value maximization, and Organizational Capabilities, *Journal of Business* 74(3), 409-431.

[20] Panzar, John, and Robert Willig, 1981, Economies of Scope, *American Economic Review* 71, 268-272.

[21] Penrose, Edith, 1959, The Theory of Growth of the Firm. Wiley, New York.

[22] Rajan, Raghuram, Servaes, Henri, and Zingales, Luigi, 2000, The Cost of Diversity: The Diversification Discount and Inefficient Investment, *Journal of Finance* 55, 35-80.

[23] Santalo, Juan, 2001, Diversification Discount, Size Discount and Organizational Capital, unpublished manuscript, University of Chicago.

[24] Scharfestein, David, 1998, The Dark Side of Internal Capital Markets II, NBER Working Paper, 6352.

[25] Scharfestein, David, and Stein, Jeremy 2000, The Dark Side of Internal Capital Markets I, *Journal of Finance* 55, 2537-2564.

[26] Schleifer, Andrei, and Robert Vishny, 1989, Managerial Entrenchment, The Case of Manager-Specific Investment, *Journal of Financial Economics* 25, 123-139.

[27] Schoar, Antoinette, 2001, Effects of Corporate Diversification on Productivity, unpublished manuscript, Massachusetts Institute of Technology.

[28] Stokey, Nancy and Lucas, Robert, 1989, Recursive Methods in Economic Dynamics, Harvard University Press.

[29] Stulz, Rene, 1990, Managerial Discretion and Optimal Financing Policies, *Journal of Financial Economics* 26, 3-27.

[30] Tirole, Jean, 1995, The Theory of Industrial Organization, MIT Press, Cambridge, MA.

[31] Villalonga, Belen, 2001, Does Diversification Cause the "Diversification Discount"?, unpublished manuscript, Harvard University.

[32] Wernerfelt, Birger and Cynthia A. Montgomery, 1988, Tobin's Q and Importance of Focus in Firm Performance, *American Economic Review* 78, 246-250

[33] Williamson, Oliver, 1975, Markets and Hierarchies: Analysis and Antitrust Implications. Free Press, New York.

A Proofs

Proof of Proposition 1. Define the operator

$$
\begin{aligned}
(Tv)(s,k,z) &= \max_{\{k',s'\}} \left\{ \pi(s',k,z) + (1-\delta)k - k' + \beta \int v(s',k',z') Q(dz',z) \right\}, \quad \text{(A1)} \\
s' &\in \begin{cases} \{s,3\}, & s = 1,2 \\ S, & s = 3 \end{cases} .
\end{aligned}
$$

Let $C(S \times K \times Z)$ be the space of all bounded and continuous functions in $S \times K \times Z$. The proof is in two steps:

1. $T : C(S \times K \times Z) \longrightarrow C(S \times K \times Z)$ (Lemma 1);

2. T is a contraction in $C(S \times K \times Z)$ (Lemma 2).

The Contraction Mapping Theorem then guarantees that there is a unique fixed point that satisfies (A1).

Lemma 1 $T : C(S \times K \times Z) \longrightarrow C(S \times K \times Z)$.

Proof. Suppose $v(s',k',z') \in C(S \times K \times Z)$. Since $Q(dz'|z)$ has the Feller property it follows from Lemma 9.5 in Stokey and Lucas (89) that

$$\int v(s',k',z') Q(dz',z) \in C(S \times K \times Z).$$

Since $\pi(s',k,z)$ is also bounded and continuous, the result follows immediately. ■

Lemma 2 T *is a contraction in* $C(S \times K \times Z)$.

Proof. The proof uses Blackwell's sufficient conditions for a contraction.

(a)Monotonicity.

Consider $v_1(s,k,z), v_2(s,k,z) \in C(S \times K \times Z)$, such that $v_1(s,k,z) \geq v_2(s,k,z)$. It follows that

$$\int v_1(s',k',z') Q(dz',z) \geq \int v_2(s',k',z') Q(dz',z),$$

and hence

$$(Tv_1)(s,k,z) \geq (Tv_2)(s,k,z).$$

(b) Discounting

Let $a \in R$ and $v(s,k,z) \in C(S \times K \times Z)$. It follows that

$$(Tv + a)(s,k,z) = v(s,k,z) + \beta a = (Tv)(s,k,z) + \beta a.$$

$\blacksquare$

Proof of Proposition 2

- $v(\cdot, k, \cdot)$ is strictly increasing in $k \in K$.

Follows from Theorem 9.7 in Stokey and Lucas (1989).

- $v(\cdot, z)$ is strictly increasing in $z \in Z$.

Follows from Theorem 9.11 in Stokey and Lucas (1989).

Proof of Lemma 3 (i) $\Pi(s,k,z) \geq 0$.

By definition

$$\pi(3,k,z) = \pi(1,\theta^* k, z) + \pi(2,(1-\theta^*)k,z) + \lambda f$$

where $\theta^* = \theta(k,z)$, is the optimal share of capital allocated to sector 1. Clearly then

$$\pi(3,k,z) \equiv \begin{cases} \pi(1,k,z) - (1-\lambda)f, & \theta^* = 1 \\ \pi(2,k,z) - (1-\lambda)f, & \theta^* = 0 \end{cases}$$

since θ^* is chosen optimally it follows that

$$\Pi(s,k,z) = \pi(3,k,z) - \pi(s,k,z) \geq 0$$

(ii) Monotonicity in k

Taking derivatives of $\pi(3,k,z)$ with respect to k we obtain

$$\frac{\partial \pi(3,k,z)}{\partial k} = \frac{\partial \pi(1,\theta^* k, z)}{\partial(\theta k)} \left(k \frac{\partial \theta^*}{\partial k} + \theta^* \right) + \frac{\partial \pi(2,(1-\theta^*)k,z)}{\partial(\theta k)} \left(-k \frac{\partial \theta^*}{\partial k} + (1-\theta^*) \right)$$

Noting that the optimal choice of θ^* implies

$$\frac{\partial \pi(1,\theta^* k, z)}{\partial(\theta k)} = \frac{\partial \pi(2,(1-\theta^*)k,z)}{\partial(\theta k)},$$

we immediately obtain

$$\frac{\partial \pi(3,k,z)}{\partial k} = \frac{\partial \pi(s',\theta^* k,z)}{\partial(\theta k)} \geq \frac{\partial \pi(s,k,z)}{\partial k}, \quad s=1,2.$$

Where the inequality follows from the fact that the profit function is strictly concave and $\theta \leq 1$. Since

$$\frac{\partial \Pi(s,k,z)}{\partial k} = \frac{\partial \pi(3,k,z)}{\partial k} - \frac{\partial \pi(s,k,z)}{\partial k} \geq 0, \quad s=1,2$$

(iii) Monotonicity in z.

Taking derivatives of $\pi(3,k,z)$ with respect to z^s and simplifying as in (ii) we obtain

$$\frac{\partial \pi(3,k,z)}{\partial z^s} = \frac{\partial \pi(s,\theta^* k,z)}{\partial z^s} + \frac{\partial \pi(\widetilde{s},(1-\theta^*)k,z)}{\partial z^s} = \frac{\partial \pi(s,\theta^* k,z)}{\partial z^s}$$

since production in sector $\widetilde{s}$ does not depend on the shock to sector i. Now, using the envelope theorem and the profits definitions (3) and (2) yields

$$\frac{\partial \pi(3,k,z)}{\partial z^s} = F(\theta^* k^s,\cdot) \leq F(k^s,\cdot) = \frac{\partial \pi(s,k,z)}{\partial z^s}.$$

and hence that

$$\frac{\partial \Pi(s,k,z)}{\partial z^s} \leq 0, \quad s=1,2.$$

Monotonicity in $z^{\widetilde{s}}$ follows immediately from the fact that $\pi(s,k,z)$ depends only on z^s.

Proof of Lemma 4 (i) $\Psi(s,z) \geq 0$.

First note that

$$\begin{aligned} v(3,k',z') &= \max\{p(1,k',z'),p(2,k',z'),p(3,k',z')\} \\ &\geq \max\{p(s'',k',z'),p(3,k',z')\} = v(s',k',z'), \\ \forall (k',z') &\in K \times Z, \forall s'' \in \{1,2\}, \end{aligned}$$

From monotonicity of $Q(\cdot)$ it follows that

$$\begin{aligned} &\int \max\{p(1,k',z'),p(2,k',z'),p(3,k',z')\}\, Q(dz',z) \\ \geq &\int \max\{p(s'',k',z'),p(3,k',z')\}\, Q(dz',z), \end{aligned}$$

Hence for any value of $z \in Z$ and any value of $k' \in K$

$$\beta \int v(3, k', z')Q(dz', z) - k' \geq \beta \int v(s', k', z')Q(dz', z) - k'.$$

Since this holds for every value of k' it follows that it holds at the maximum and $\Psi(s, z) \geq 0$.

(ii) $\Psi(s, z)$ is decreasing in z^s and increasing in $z^{\widetilde{s}}$, $s \neq \widetilde{s}$.

Suppose $z^s >> z^{\widetilde{s}}$. Then

$$p(s, k, z) >> p(\widetilde{s}, k, z),$$

and consequently

$$v(3, k, z) \approx \max \{p(s, k, z), p(3, k, z)\}.$$

Given the monotonicity of $Q(\cdot)$ it follows that:

$$\int v(3, k', z')Q(dz', z) \approx \int \max \{p(s, k', z'), p(3, k', z')\} Q(dz', z) = \int v(s, k', z')Q(dz', z),$$

and, therefore, $\Psi(s, z) = 0$.

Now suppose that the opposite is true, i.e. $z^s << z^{\widetilde{s}}$. In that case

$$v(3, k, z) \approx \max \{p(\widetilde{s}, k, z), p(3, k, z)\}$$

and

$$\int v(3, k', z')Q(dz', z) \approx \int \max \{p(\widetilde{s}, k', z'), p(3, k', z')\} Q(dz', z) > \int v(s, k', z')Q(dz', z).$$

which implies that $\Psi(s, z) > 0$. It follows from continuity of both $v(\cdot)$ and $Q(\cdot)$ that $\Psi(s, z)$ must fall with z^s.

An identical argument can be constructed to establish that $\Psi(s, z)$ increases with $z^{\widetilde{s}}$.

Proof of Lemma 13 Note that given Assumption 3, $p(s'', \cdot) \geq 0$. Hence $\max_{k'} \left\{ \int p(s'', \cdot)Q(dz', z) - k' \right\} \geq 0$. Depending on the sectoral choice of the firm next period, there are three possible outcomes for the maximizations:

(a) $p(3, \cdot) \geq \max \{p(2, \cdot), p(1, \cdot)\}$, and

$$\begin{aligned} & \max_{k'} \left\{ \int p(3, \cdot)Q(dz', z) - k' \right\} \\ \leq \; & \max_{k'} \left\{ \int p(3, \cdot)Q(dz', z) - k' \right\} + \max_{k'} \left\{ \int p(3, \cdot)Q(dz', z) - k' \right\}; \end{aligned}$$

(b) $p(3,\cdot) \leq p(2,\cdot) \wedge p(3,\cdot) \leq p(1,\cdot)$, then

$$\begin{aligned} & \max_{k'} \left\{ \int \max\left\{ p(2,\cdot), p(1,\cdot) \right\} Q(dz', z) - k' \right\} \\ \leq \ & \max_{k'} \left\{ \int p(1,\cdot) Q(dz', z) - k' \right\} + \max_{k'} \left\{ \int p(2,\cdot) Q(dz', z) - k' \right\}; \end{aligned}$$

(c) $p(s_1'',\cdot) \geq p(3,\cdot) \geq p(s_2'',\cdot)$, $s_1'', s_2'' \in \{1,2\}$, and

$$\begin{aligned} & \max_{k'} \left\{ \int p(s_1'',\cdot) Q(dz', z) - k' \right\} \\ \leq \ & \max_{k'} \left\{ \int p(s_1'',\cdot) Q(dz', z) - k' \right\} + \max_{k'} \left\{ \int p(3,\cdot) Q(dz', z) - k' \right\}; \end{aligned}$$

Hence, our result holds in all cases.