



American Finance Association

Optimal Diversification: Reconciling Theory and Evidence

Author(s): Joao Gomes and Dmitry Livdan

Source: *The Journal of Finance*, Vol. 59, No. 2 (Apr., 2004), pp. 507-535

Published by: Blackwell Publishing for the American Finance Association

Stable URL: <http://www.jstor.org/stable/3694906>

Accessed: 23/11/2009 13:03

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Blackwell Publishing and American Finance Association are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance*.

<http://www.jstor.org>

Optimal Diversification: Reconciling Theory and Evidence

JOAO GOMES and DMITRY LIVDAN*

ABSTRACT

In this paper we show that the main empirical findings about firm diversification and performance are consistent with the maximization of shareholder value. In our model, diversification allows a firm to explore better productive opportunities while taking advantage of synergies. By explicitly linking the diversification strategies of the firm to differences in size and productivity, our model provides a natural laboratory to investigate several aspects of the relationship between diversification and performance. Specifically, we show that our model can rationalize the evidence on the diversification discount (Lang and Stulz (1994)) and the documented relation between diversification and productivity (Schoar (2002)).

EMPIRICAL WORK ON FIRM DIVERSIFICATION has often been interpreted as supporting the view that conglomerates are inefficient. Findings such as the fact that conglomerates trade at a discount, relative to a portfolio of comparable stand-alone firms, have led researchers to believe that diversification destroys value.¹ Popular explanations for this “diversification discount” have generally emphasized the agency and behavioral problems associated with the existence of conglomerates.² Unfortunately, this view of diversification creates at least two difficulties for researchers. First, while addressing the effects of diversification on performance, agency models often fail to answer the more fundamental economic question of why diversified firms exist at all, as diversification is often ex-ante inefficient. Second, the empirical predictions of these agency-based models are usually very hard to quantify and thus quite difficult to test. As a consequence, direct evidence supporting this agency view is

*The Wharton School, University of Pennsylvania and Bauer College of Business, University of Houston. We are grateful to Andrew Abel, Michael Brandt, Domenico Cuoco, Jan Eberly, Simon Gervais, Armando Gomes, Francisco Gomes, Gary Gorton, John Graham, Skander van den Heuvel, Rich Kihlstrom, Leonid Kogan, Jan Mahrt-Smith, Vojislav Maksimovic, Andrew Metrick, Gordon Phillips, Tom Sargent, Jeremy Stein, and an anonymous referee, as well as to seminar participants at Kellogg, Wharton, Penn State, the 2002 SED and Econometric Society Meetings, and the 2003 AFA meetings. Financial support from the Rodney L. White Center for Financial Research is gratefully acknowledged. This paper combines our earlier papers “Optimal Diversification” and “The Performance of Optimally Diversified Firms: Reconciling Theory and Evidence.”

¹ See Wernerfelt and Montgomery (1988), Lang and Stulz (1994), Berger and Ofek (1995), Rajan, Servaes, and Zingales (2000), Whited (2001), Lamont and Polk (2001, 2002), among others.

² See Jensen (1986), Amihud and Lev (1981), Jensen and Murphy (1990), Shleifer and Vishny (1989), and Stulz (1990), Denis, Denis, and Sarin (1997), and Scharfstein and Stein (2000) among others.

quite limited. Instead, support typically comes from the perceived failures of competing theories.

In this paper we show that the main empirical regularities about firm diversification are broadly consistent with a neoclassical view of efficient firm diversification. In our model, firms diversify for two reasons. First, diversification allows firms to take advantage of economies of scope by eliminating redundancies across different activities and lowering fixed costs of production. Second, diversification allows a mature, slow-growing firm to explore attractive new productive opportunities. We formalize this concept by assuming that production activities exhibit decreasing returns to scale. As scale grows, returns decrease, eventually leading the firm to search for profit opportunities in new activities.

In contrast to standard agency arguments, our model's structure provides a natural environment to investigate *quantitatively* the role of firm diversification on performance. Since the model generates an artificial cross-sectional distribution of firms, we are able to directly compare our results with the available empirical evidence.

We have two main sets of findings. First, the model predicts that diversified firms have, on average, a lower value of Tobin's q than focused firms, as documented by Lang and Stulz (1994). This happens despite the fact that diversification is optimal and there is no source of inefficiency in our model. The intuition, however, is simple. In our model, firms diversify only when they become relatively unproductive in their current activities. It is this *endogenous selection* mechanism that accounts for the lower valuation of diversified firms. Second, because our model explicitly links productivity with corporate diversification, we can also address recent evidence about the effects of diversification on productivity (Schoar (2002)). We find that, just as in the data, our model predicts that firms following diversification strategies also experience empirically plausible productivity losses.

This emphasis on the importance of firm selection in accounting for the performance of conglomerates effectively presents a theoretical foundation for the recent empirical findings by Chevalier (2001), Villalonga (2001), and Campa and Kedia (2002). Although their exact sources and methodologies differ, all of these papers are part of a growing empirical literature suggesting that sample selection accounts for most, if not all, of the ex-post differences between conglomerates and specialized firms.³

More broadly, our work is also part of a recent strand of literature that emphasizes a view of conglomerates as profit-maximizing firms. Nevertheless, all of them still assume that diversification reduces firm value; while conglomerates allocate resources efficiently (profit maximization), they are still endowed with lower profit opportunities than a specialized firm (diversification is value-reducing).

³ Although we do not address their findings, recent work by Whited (2001), Graham, Lemmon, and Wolf (2002) also suggests that the diversification discount is not caused by inefficient conglomerates.

For example, 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 eventually specializes. Bernardo and Chowdhry (2002) explain the diversification discount by assuming that specialized firms have growth options allowing them to diversify in the future. Conglomerates, on the other hand, are firms that have exercised these options and are thus less valuable to investors.

Closest to our model is the work of Maksimovic and Phillips (2002), who first formalize the idea that diversification decisions can be understood as the optimal response of firms to industry or sectoral shocks. Using a static linear quadratic model, they show that firms will become conglomerates only when they face similar profit opportunities across sectors. Specialized firms, on the other hand, are usually much more productive in their chosen activities. Their paper also provides strong supporting evidence for this view. In their model conglomerates are valued at a premium relative to small specialized firms and at a discount relative to large specialized firms. However, they also assume that firms must incur extra costs when they produce in more than one industry.

Thus, while our dynamic environment does incorporate features from these models, our analysis differs in one crucial way. In our model conglomerates are not assumed to be *ex ante* less efficient. Thus our model generates a diversification discount endogenously, an explanation that seems consistent with recent empirical evidence, and questions the common interpretation of the diversification discount as evidence of inefficiency. This endogenous relation between productivity and diversification is also present in Maksimovic and Phillips (2002) and Inderst and Muller (2003). Our dynamic setting, however, allows us to also account for the interaction between productivity and firm size as joint determinants of the decision to diversify. As we will see, this allows us to better account for the existing empirical evidence.

In addition to this key distinction, our model also provides a unified and consistent explanation for much of the empirical evidence by endogenously linking productivity, size, and valuations to diversification strategies. Finally, our approach relies on the detailed quantitative evaluation of an artificial panel of firms generated by solving a fully specified general equilibrium environment. Thus our framework is able to produce a well-defined cross-sectional distribution of firms that provides both a reasonable description of the data and a natural ground for examining the quantitative implications of our model.

Finally, our work also offers a useful framework to study the natural boundaries of the firm in the context of a neoclassical environment. While our model is silent about the exact micro-foundations for the interactions between (and within) firms (e.g., internal capital markets, incomplete contracts, and power relationships within firms), it provides something of a reduced form approach that is well suited for detailed empirical study, a serious difficulty in this field of research.

The rest of this paper is organized as follows. Section I details the basic economic environment and discusses our main assumptions. Section II provides

a quantitative evaluation of our model and establishes its main empirical implications. Section III concludes.

I. Model

A pattern that emerges from much of the relevant empirical evidence is one of substantial firm heterogeneity across a number of different characteristics such as firm size and growth, as well as investment and diversification strategies. It is therefore crucial that our framework is consistent with this evidence and is thus able to produce a well-defined cross-sectional distribution of firms that provides a reasonable description of the data.

Our theoretical approach is then based on an industry equilibrium environment with heterogeneous firms, along the lines of Hopenhayn (1992) and Gomes (2001). Our economy consists of two sectors: households and firms. The core of the analysis is our description of the production sector, where a large number of firms is engaged in the production of the consumption good. The role of households is limited and is summarized by a single representative household making optimal consumption and portfolio decisions.

A. Firms

The production side of the economy consists of a large number of firms and two separate industries or sectors. While the model can be augmented to include more sectors, this would make the analysis unnecessarily complicated. Empirically, the effects of diversification on performance are most notable when firms first expand from one to two segments, with additional expansions having only marginal effects on performance (Lang and Stulz (1994)).

A.1. Description

We assume that time is discrete and the horizon is infinite. In each time period t , a firm can either be focused in sector $s_t = 1, 2$ or operate in both sectors simultaneously, in which case we will say that a firm is diversified and set $s_t = 1 + 2 = 3$. We assume that sectoral mobility is costly so that specialized firms cannot simply move all resources from sector one to sector two (say). Formally, we assume that

$$s_t \in \begin{cases} \{s_{t-1}, 3\}, & s_{t-1} = 1, 2 \\ \{1, 2, 3\}, & s_{t-1} = 3. \end{cases} \quad (1)$$

In other words, a firm that has previously been focused in sector s can only choose to remain in sector s ($s_t = s_{t-1}$), or to expand to both sectors ($s_t = 3$). Diversified firms, however, face no restrictions. They can either remain diversified, or they can refocus on a single industry. This costly mobility ensures that

a firm must diversify before focusing on entirely new activities, a pattern that is consistent with the data.⁴

The outcome of production in sector s , during period t , is the final good y_t^s . For simplicity, we assume that the goods are perfect substitutes so that the relative price between y_t^1 and y_t^2 is always equal to one. Production in either sector requires two inputs: capital or productive capacity, k_t , and labor, l_t , and is subject to a technology shock z_t^s . Labor is hired at the competitive wage rate $W_t > 0$, but capacity is owned by the firm. Production possibilities for an individual firm operating in sector s are described by a Cobb–Douglas production function

$$y_t^s = e^{z_t^s} k_t^{\alpha_k} l_t^{\alpha_l}, \quad 0 < \alpha_k + \alpha_l < 1, \quad (2)$$

where α_k and α_l are the output elasticities of capital and labor, respectively. The restrictions on these coefficients guarantee that production in each sector exhibits decreasing returns to scale, so that returns fall as the firm grows.

Productivity levels are firm specific and cannot be traded. We assume that productivity in each sector s follows a simple AR(1) process

$$z_t^s = \rho z_{t-1}^s + \varepsilon_t^s, \quad (3)$$

where each ε_t^s is a normal random variable with mean zero and variance σ^2 . For simplicity we also assume that there is no cross-correlation between the shocks in the two sectors. To save on notation we also define the productivity vector $z_t = (z_t^1, z_t^2)$.

Finally, total firm capacity is described by the law-of-motion

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

where i_t denotes gross investment spending, and δ is the depreciation rate of capital. Thus, new investment, i_t , becomes productive only at the beginning of the next period.

The timing of the decisions is illustrated in Figure 1.

Every firm arrives at period t with a pre-chosen level of capacity k_t . Before any activity takes place, the firm observes the (firm-specific) vector of productivity levels in both sectors, z_t . With this information at hand, each firm makes the following choices during the period t :

- the optimal sectoral decision for the current period, s_t , by choosing whether to operate one ($s_t = 1$ or 2) or both ($s_t = 3$) production units in period t ;
- the optimal allocation of capital and labor across its activities;

⁴ This assumption has little impact on our numerical analysis. It will, however, make it easier to construct Figure 2 and to gain some intuition about our results. A more realistic model would allow for some form of costly switching, possibly as a consequence of learning or irreversible investment, but this would greatly complicate our analysis.

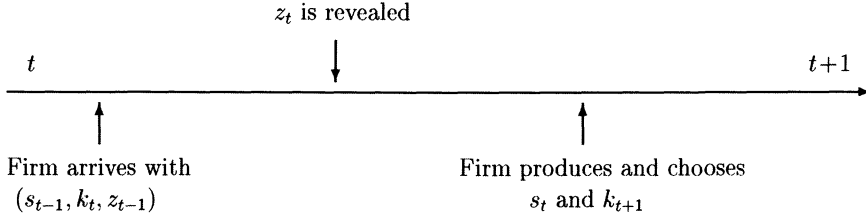


Figure 1. Timing of events. This figure illustrates the timing of events. Every firm arrives at period t with a pre-specified level of capacity k_t . Before any activity takes place, the firm observes a vector of firm-specific productivity shocks in both sectors, z_t . With this information in hand, the firm chooses in the following order: its sectoral allocation, s_t , the optimal allocation of capital across its activities, and how much capacity to install at the beginning of the next period, k_{t+1} .

- how much to invest for the future, i_t , and, as a consequence, the total amount of capacity to install at the beginning of the next period, k_{t+1} .

A firm that chooses to focus its activities in sector s_t alone generates the following profits during period t

$$\pi(s_t, k_t, z_t; W_t) = \max_{l_t} \{e^{z_t^s} k_t^{\alpha_k} l_t^{\alpha_l} - W_t l_t - f\}, \quad s_t = 1, 2, \quad (5)$$

where $f \geq 0$ is a fixed cost of production that must be paid if the firm is active in sector s .⁵ Conversely, if the firm chooses to be diversified (so that $s_t = 3$), profits are described by

$$\pi(3, k_t, z_t; W_t) = \max_{l_t, \theta_t} \{e^{z_t^1} (\theta_t k_t)^{\alpha_k} (\theta_t l_t)^{\alpha_l} + e^{z_t^2} ((1 - \theta_t) k_t)^{\alpha_k} ((1 - \theta_t) l_t)^{\alpha_l} - W_t l_t - (2 - \lambda)f\}, \quad (6)$$

$$s.t. \quad 0 \leq \theta_t \leq 1, \quad (7)$$

where θ_t denotes the fraction of resources (capital and labor) that the diversified firm allocates to sector one in period t .⁶ Because diversified firms operate in both sectors, they face larger fixed costs of production. However, equation (6) embeds our assumption that they can eliminate redundancies and thus save a fraction $\lambda/2$ of the combined costs in each sector. Thus, a conglomerate pays only fixed costs in the amount $(2 - \lambda)f$.

The solution to these static optimization problems yields optimal decision rules for total firm employment, $l_t = \mathbf{l}(s_t, k_t, z_t; W_t)$, the size of each segment, $\theta_t = \theta(s_t, k_t, z_t; W_t)$, as well as total production, $y_t = \mathbf{y}(s_t, k_t, z_t; W_t)$.

⁵ Fixed costs guarantee a minimum scale of production, thus forcing a firm to stay focused, unless outside opportunities are sufficiently attractive. As we show below, without fixed costs a firm will always be diversified.

⁶ Since both output elasticities and prices do not differ across sectors, capital-labor ratios must also be identical. It follows that the conglomerate must allocate the same share of capital and labor inputs to each sector.

A.2. Discussion

Our environment is constructed to incorporate the main incentives for the creation of conglomerates identified by the literature on firm diversification. Somewhat loosely our model emphasizes some of the most popular advantages of firm diversification: “Synergies” and the exploration of “free” cash flows.

Synergies are created through the elimination of redundancies across business lines, such as overhead. In our model, this feature is captured by the savings parameter $\lambda \geq 0$. Such dilution of costs generates a form of economies of scope and creates a benefit to firm diversification that cannot be replicated by shareholders. It is this key advantage that separates our work from the existing literature. In our model, conglomerates not only operate efficiently, but they also create value for investors. As a result, the resulting diversification discount in our model is driven entirely by the endogenous nature of the diversification decision.

In addition to this key feature, our model also assumes the existence of decreasing returns to scale in each sector. This assumption generates something like a “free cash flow” effect. As the firm grows in size, marginal productivities fall and it becomes unprofitable for the firm to invest additional resources in ongoing activities. Instead, the firm can better use resources by exploring new production possibilities. Thus, diversification is more likely to be optimal for large firms, since it enables them to overcome the decreasing returns nature of the single-sector technology. This feature is also consistent with the empirical observation that large firms are much more likely to become diversified.

Decreasing returns to scale are also used by Santalo (2001) and Maksimovic and Phillips (2002). Santalo constructs a model where the diversification discount is attributed to the “size discount” induced by decreasing returns. However, much of the empirical evidence suggests that decreasing returns alone are not sufficient, since the diversification discount seems to survive after one controls for the size of the firm. Maksimovic and Phillips use a linear quadratic example to show how decreasing returns to scale can provide a natural bound to the size of the firm and thus create an incentive to diversification. Because they do not have fixed costs in their model, however, firms would never find it optimal to be focused. To overcome this, they instead must assume that conglomerates have higher production costs than those of two focused firms combined.

In addition to these core advantages, conglomerates also benefit from two additional features of our environment. They have more options than stand-alone firms (the mobility restriction (1)). Although this is not a crucial feature of our model, it stands in contrast to Bernardo and Chowdhry (2002), who rationalize the diversification discount by assuming that focused firms have more options than conglomerates.

Finally, since the productivity shocks z_t^1 and z_t^2 are not perfectly correlated (as in equation (3) above), diversification allows a firm both to explore alternative profit opportunities and to lower exposure to cash flow risk. This idea of conglomerates as the optimal response of firms to sectoral variations in profit opportunities was first introduced to this literature by Maksimovic and Phillips (2002). However, in the absence of trading frictions, this feature is not valued by

investors in general equilibrium since it can be easily replicated by a portfolio of stand-alone firms.

Thus, in our model, production is more efficient and resources are saved when operations are combined into a conglomerate. Hence, unlike much of the existing literature on corporate diversification, our model captures some of the most plausible benefits to corporate diversification while abstracting from any of its potential drawbacks, such as those induced by agency or behavioral problems.

Emphasizing these advantages of the conglomerates ensures that our model does not deliver a diversification discount “by assumption.” Since conglomerates generally have more resources and better opportunities in our model, their low valuation can only be an outcome of the *endogenous* self-selection, and not an obvious consequence of assuming that focused firms are, a priori, better. As a number of recent studies suggest, this explanation seems to be consistent with the available evidence.

A.3. Optimality

Let (s, k, z) denote the state for a firm that was active in sector s in period $t - 1$, has k units of installed capacity at the beginning of period t , and faces a vector of productivity shocks z . The optimal behavior of this firm can be summarized by the value function $v(s, k, z; W)$, that solves the dynamic programming problem:

$$v(s, k, z; W) = \max_{k', s'} \left\{ \pi(s', k, z; W) + (1 - \delta)k - k' + \beta \int v(s', k', z'; W') N(dz' | z) \right\} \quad (8)$$

subject to equation (1).⁷ Here $0 < \beta < 1$ is the intertemporal discount factor and $N(dz' | z)$ is the cumulative (Gaussian) distribution of z' , conditional on z . Note that current cash flows (dividends) are given by current profits, $\pi(\cdot)$, net of investment spending, i , which is described by equation (4). Proposition 1 establishes the existence of a unique function $v(s, k, z)$, that satisfies equation (8), and lists some of its basic properties.

PROPOSITION 1: *There exists a unique function $v(s, k, z)$ that solves the dynamic program (8). Moreover, this function is (i) continuous and (ii) increasing in both k and z .*

See Appendix A for the proof.

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. Finally, the solution to the dynamic programming problem (8) also produces a set of policy functions,

⁷ We use the convention s', k', z' , etc., to denote the value of the state variables that are relevant at the beginning of the next period.

$\mathbf{k}(s, k, z; W)$ and $\mathbf{s}(s, k, z; W)$, associated with the optimal accumulation of capital and the sectoral choices of the firm. It is straightforward to show that all these functions are well defined.

A.4. The Decision to Diversify

Before exploring the quantitative implications of the model, it is useful to study some of the inner workings of our model. Accordingly, this section attempts to shed some light on the optimal diversification decision of an individual firm.

The optimal industrial decision, $s' = \mathbf{s}(s, k, z)$, can be computed as follows. First, define the function

$$p(s', k, z) \equiv \pi(s', k, z) + (1 - \delta)k + \max_{k'} \left\{ \beta \int v(s', k', z') N(dz' | z) - k' \right\} \quad (9)$$

as the value of the firm, *conditional* on having adopted sectoral decision s' in the current period. Since focused firms are not allowed to simply switch sectors, a firm that was previously specialized in sector $s \in \{1, 2\}$, finds corporate diversification optimal if, and only if

$$p(3, k, z) \geq p(s', k, z) \mid_{s'=s} = p(s, k, z). \quad (10)$$

Similarly, a firm that was diversified in the previous period ($s = 3$) will choose to remain diversified if

$$p(3, k, z) \geq \max\{p(1, k, z), p(2, k, z)\}. \quad (11)$$

However, it is probably more useful to represent this decision on the space of state variables. Proposition 2 shows how this can be done, by defining a “diversification threshold,” $\hat{k}$, that divides the state space into different regions according to the optimal strategy of the firm.⁸ Proposition 2 establishes the key properties of this threshold.

PROPOSITION 2: *The optimal diversification decision can be characterized by the unique threshold value:*

$$\hat{k}(s, z) = \arg \min_k \{\mathbf{s}(s, k, z) = 3\}, \quad \forall (s, z) \in S \times Z. \quad (12)$$

Moreover, $\hat{k}(s, z)$, is: (i) increasing in z^s and, (ii) decreasing in $z^{\bar{s}}$, $s \neq \bar{s}$.

See Appendix A for the proof.

⁸ Formally, this threshold is a separating hyperplane in the four-dimensional space of state variables.

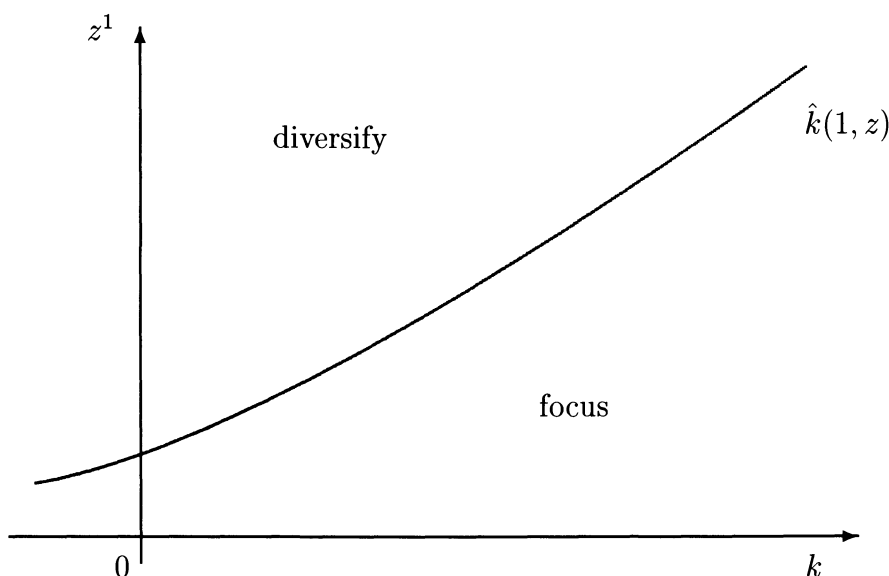


Figure 2. The diversification threshold. This figure illustrates the shape of the optimal sectoral decision, $s(1, k, z)$, for a firm that was previously focused in sector one. The horizontal axis shows capacity, k , and the vertical axis shows the level of productivity in on-going activities, z^1 . Since the firm was previously focused in sector one, it has only two choices: It can either remain in sector one in the current period, or it can diversify and operate in both sectors simultaneously. The figure shows the contour line of the optimal sectoral decision, holding the level of productivity in the other sector, z^2 , fixed. Points along this line correspond to combinations of productivity, z , and size, k , for which the firm is indifferent between focusing and diversifying.

Figure 2 illustrates these results by showing the shape of the optimal sectoral decision for a firm previously focused in sector one, $s(1, k, z)$. The figure depicts the diversification threshold, holding the level of z^2 fixed. Remember that this firm can only choose to remain in sector one, or to diversify. By definition, points along this line correspond to combinations of productivity, z , and size, k , for which the firm is indifferent between focusing and diversifying.

The positive slope of $\hat{k}(1, z)$ implies that firms of the same size 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. Similarly, holding productivity constant, diversification is more likely for large firms, a consequence of decreasing returns to scale. It is this *endogenous* selection feature of our model that drives several of our quantitative results below, and in particular, our findings of a diversification discount in the cross-section of firms.⁹ Thus, the model formalizes the argument proposed in several

⁹This endogenous relation between productivity and diversification is also present in Maksimovic and Phillips (2002) and Inderst and Muller (2003). Our dynamic setting, however, allows us to also account for the interaction between productivity and firm size as joint determinants of the decision to diversify. As we will see, this allows us to better account for the existing empirical evidence.

empirical studies (see Chevalier (2001), Villalonga (2001), Graham, Lemmon, and Wolf (2002), and Campa and Kedia (2002)) that conglomerates are not simply a random subsample of the cross-sectional distribution of firms. Instead, because the decision to diversify is endogenous, it is associated with ex-ante differences in firm-specific features such as productivity and size. These ex-ante features account for the findings about ex-post performance and valuation of conglomerates.

Corollaries 1 and 2 establish two additional properties of the optimal industrial strategy $\mathbf{s}(s, k, z)$. Corollary 1 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 2 shows that if synergies are sufficiently large, there is never an incentive for the firm to be focused.

COROLLARY 1: *In the absence of fixed costs ($f = 0$), diversification is always optimal.*

See Appendix A for the proof.

COROLLARY 2: *Suppose $f > 0$. Diversification is the optimal corporate strategy if $\lambda \geq 1$, i.e., synergies are sufficiently large.*

See Appendix A for the proof.

B. Aggregation and Equilibrium

To provide a detailed evaluation of the implications of our model, we need to construct an artificial panel of firms that can then be used to examine the available empirical evidence. We can do this by aggregating the individual decisions of every firm in the economy and computing the equilibrium in our model. Since each firm can be described by the (s, k, z) , the cross-sectional distribution of firms is completely summarized by a measure, $\mu(s, k, z)$, defined over this state space. The law of motion for μ is given by

$$\mu'(s', k', z') = \int \mathbf{1}_{\{k' = \mathbf{k}(s, k, z; W)\}} \times \mathbf{1}_{\{s' = \mathbf{s}(s, k, z; W)\}} N(z' | dz) \mu(ds, dk, z), \quad (13)$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function that equals one if the argument is satisfied and zero otherwise. Intuitively, next period's cross-sectional distribution of firms is determined by combining the exogenous transition probabilities implied by $N(\cdot)$ with the endogenous ones, prescribed by the optimal policies for capacity, $\mathbf{k}(s, k, z)$, and sectoral choices, $\mathbf{s}(s, k, z)$. For empirical purposes we are interested in the properties of a *stationary* equilibrium where this distribution does not depend on initial conditions, so that $\mu' = \mu$.

To close the model, we must offer a description of market demand for the final goods produced, as well as the supply of labor input. While it is easy to provide reduced form expression for these functions, it is also straightforward

to show how this can be done in general equilibrium by adding a very stylized description of household/shareholder behavior. Specifically, we summarize the household sector with a single representative agent deriving utility from leisure, L , and consumption, C , and income from wages, W , and dividends, D . Without aggregate uncertainty, all aggregate quantities and prices are constant and the consumer problem collapses to the static representation

$$\max_{C,L} U = \ln(C - AL) \quad (14)$$

$$s.t. \quad C = WL + D. \quad (15)$$

The optimality conditions for this problem yield a demand for final goods given by $C = C(\mu; W, D)$ and an infinitely elastic labor supply, which pins down the wage rate at $W = A$.¹⁰

A stationary equilibrium for this economy is characterized by the following definition.

Definition 1 (stationary equilibrium): A stationary competitive equilibrium is: (i) a set of optimal policies $\mathbf{k}(\cdot)$, $\mathbf{l}(\cdot)$, $\pi(\cdot)$, and $\mathbf{s}(\cdot)$ and a value function $v(\cdot)$ for each firm; (ii) an optimal allocation rule $C(\mu; W)$ for the representative household; (iii) a wage rate W ; and (iv) a stationary measure μ of firms such that

$$C(\mu; A) = A \int \mathbf{l}(s, k, z; W) \mu(ds, dk, dz) + \mathbf{D}(\mu; A). \quad (16)$$

Equation (16) also summarizes labor market equilibrium, by imposing $W = A$. It also uses the fact that aggregate dividends are given by¹¹

$$\begin{aligned} \mathbf{D}(\mu; A) = & \int \pi(s, k, z; A) \mu(ds, dk, dz) \\ & - \int (\mathbf{k}(s, k, z; A) - (1 - \delta)k) \mu(ds, dk, dz). \end{aligned} \quad (17)$$

Given our assumptions, establishing the existence of a stationary competitive equilibrium is immediate.¹² Although the definition seems abstract and its computation is nontrivial, this equilibrium concept is the key to our analysis. It delivers a nondegenerate cross-sectional distribution of firms, μ , which provides us with an artificial data set of firms of different size, productivities, and more importantly, diversification strategies. With this information at hand, we are ready to address the key empirical findings in this area.

¹⁰ This problem is equivalent to that of a shareholder investing in the stocks of each firm (Gomes (2001)).

¹¹ Note that $\pi(s', k, z; W) = \pi(\mathbf{s}(s, k, z), k, z; W) = \pi(s, k, z; W)$. Similarly, for $\mathbf{l}(s, k, z; W) = \mathbf{l}(s', k, z; W)$.

¹² The proof follows the arguments provided in Hopenhayn (1992) and Gomes (2001).

II. Quantitative Results

Computing the stationary equilibrium involves two steps. First, we must specify parameter values. These must be selected to be consistent with either long-run properties of the data (unconditional first moments) or with prior empirical evidence. Second, we develop and implement a numerical algorithm capable of approximating the stationary equilibrium up to an arbitrarily small error. Appendix B describes this procedure in detail. With the equilibrium computed, we focus on two key empirical issues. Section II.B investigates the model's implications for the so-called "diversification discount," by comparing our predictions with the results in Lang and Stulz (1994). Since our model implies that diversification is driven by productivity differentials, it is important to investigate its predictions for the relation between firm diversification and productivity. Section II.C explores this issue by comparing our results with the empirical evidence in Schoar (2002).

A. Calibration and Summary Statistics

Since most data is available at an annual frequency, we assume that a time period in the model corresponds to 1 year. The calibration exercise is divided in two parts. First, we use independent evidence on the degree of returns to scale (Burnside (1996)) to set the output elasticities $\alpha_l = 0.65$ and $\alpha_k = 0.3$. The rate of depreciation in the capital stock is set to 0.1, a value close to that found in the data by Gomes (2001).

The four remaining parameters, f , λ , σ , and ρ , cannot be individually identified from the available data. Instead, they are chosen so that the model is able to approximate the *unconditional* moments on the panel studied by Lang and Stulz (1994) for COMPUSTAT. Since the main stylized facts are, in effect, *conditional* moments, or regressions, from this panel, this seems appropriate. Accordingly, we select these parameters so that the model approximates the cross-sectional mean and dispersion of Tobin's q , the fraction of diversified firms in the sample, and the average level of q for conglomerates.¹³

Table I summarizes our calibration procedure while Table II compares the key summary statistics generated by the stationary equilibrium of the model with those of the COMPUSTAT data set used by Lang and Stulz (1994). Although our model calibration does not reproduce these four statistics exactly, the artificial sample is reasonably similar to its empirical counterpart, particularly in terms of cross-sectional dispersion and the relative weight of conglomerates in the sample, the two crucial elements for statistical inference.

B. Diversification Discount

Most empirical studies on the efficiency of conglomerates examine the relation between diversification and the value of the firm, usually measured by

¹³ The preference parameters are not important and we simply use $\beta = 1/1.065$ and $A = W = 0.5$.

Table I
Parameter Choices

This table reports our parameter choices. The time period is 1 year. Output elasticities, α_k and α_l , are set using evidence from Burnside (1996). The rate of depreciation for the capital stock, δ , is set close to the value found by Gomes (2001). The four remaining parameters, f , λ , σ , and ρ , are chosen so that the model approximates four unconditional moments from the COMPUSTAT panel studied by Lang and Stulz (1994). The moments are the mean and standard deviation of Tobin's q , the number (percentage) of diversified firms in the sample, and the average level of Tobin's q for conglomerates.

Parameter	Benchmark Value
Technology	
α_k	0.3
α_l	0.65
δ	0.1
f	0.002
λ	0.6
Shocks	
σ	0.025
ρ	0.95

Table II
Summary Statistics

This table compares the summary statistics generated by the stationary equilibrium of the model, given the parameter choices in Table 1, with those of the COMPUSTAT panel studied by Lang and Stulz (1994) and reported in Table 1 of their paper.

Statistics	Data	Model
Fraction focused firms	0.40	0.33
Tobin's q		
Average	1.11	1.87
Standard deviation	1.22	1.11
Average (conglomerates)	0.91	1.56

Tobin's (average) q .¹⁴ Specifically, this is often done by estimating linear reduced form equations

$$q_{it} = b_0 + b_1 DIV_{it} + b_2 \ln(k_{it}) + \xi_{it}, \quad (18)$$

where q_{it} is the value of Tobin's q for firm i at the beginning of period t , k_{it} is the beginning of period size of the firm, and DIV_{it} is a dummy variable that takes a value of one if the firm is diversified in period t and is zero otherwise.

¹⁴ The use of Tobin's q to measure the value of the firm is theoretically correct even in the presence of measurement error.

Table III
The Diversification Discount

This table reports the results of estimating the following regression:

$$q_{it} = b_0 + b_1 DIV_{it} + b_2 \ln(k_{it}) + \xi_{it},$$

on our artificial panel of firms. Here q_{it} is the value of Tobin's q for firm i at the beginning of period t , k_{it} is the beginning of period size of the firm, and DIV_{it} is a dummy variable that takes the value 1 if the firm is diversified in period t and is zero otherwise. The results of this estimation are then compared with the empirical findings from Table 6 in Lang and Stulz (1994). In all cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics. The table also reports our findings for the subset of firms for which the value of q is below five, and compares those with the results in Lang and Stulz.

Variable	All Firms		$q < 5$	
	Data	Model	Data	Model
DIV	-0.34	-0.20	-0.29	-0.07
(t -stat)	(-3.77)	(-5.39)	(-4.53)	(-3.71)
$\log(k)$	-0.12	-0.70	-0.13	-0.31
(t -stat)	(-3.48)	(-5.26)	(-5.22)	(-5.29)

In the context of our model, it is straightforward to estimate equation (18) for our artificial panel of firms by defining the variables

$$q \equiv \frac{p(s', k, z)}{k}, \quad (19)$$

and

$$DIV \equiv \begin{cases} 1, & \text{if } s' = 3 \\ 0, & \text{else} \end{cases}, \quad (20)$$

where $p(s', k, z)$ denotes the value of the firm of size k that chooses to operate in sector s in period t .

Table III compares the results of estimating (18) in our model with the empirical findings in Lang and Stulz (1994). In all cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics. As in Lang and Stulz, Table III reports results for both the full panel and a subset that includes only those firms with a value of q below five.¹⁵

Overall the model performs very well. As in the data, we consistently find that diversified firms are discounted, that this discount is statistically significant, and that this is only partially accounted for by differences in firm size (the coefficient on $\ln(k)$). Moreover, the model also predicts a diversification discount that is *quantitatively* similar to that found in the real data.

¹⁵ When possible we focus on the numbers reported by Lang and Stulz (1994) for "industry-adjusted" q 's, since these control for the fact that diversified firms are generally concentrated in low q industries.

Looking only at the subset of firms with a value of q below five shows that the observed diversification discount is not due to a small number of outliers. Table III confirms that in the model, as in the data, eliminating outliers does decrease the discount's magnitude but it does not eliminate it. Although smaller, the coefficient on the diversification dummy is significant, both statistically and economically.

Thus, despite the fact that conglomerates operate efficiently and that diversification clearly adds value to the firm, our model is able to rationalize the documented diversification "discount." Moreover, this discount also seems to possess the same robustness properties that are observed in the actual data. Since diversification is optimal, however, the explanation cannot be that conglomerates destroy value. Instead, the success of the model hinges on the endogenous selection mechanism identified in Section A.4.

It is important to note that our results accord with the view that conglomerates are indeed less efficient firms. Crucially, however, they are not *inefficient*. In particular, and as long as $\lambda > 0$, separation of their units destroys shareholder value.

Finally, the exact magnitude of the discount depends on synergies created by the conglomerate, measured by the parameter λ . Indeed, if these synergies are too large, the discount may disappear altogether. We view this dependence as an important strength of the model and a useful direction for future research. For instance, allowing λ to vary across firms could rationalize recent evidence suggesting that the magnitude of the discount seems to vary with the level of synergies created by diversification (see, e.g., Chevalier (2001)).

B.1. Source of the Diversification Discount

Following Lang and Stulz (1994), Tables IV and V attempt to shed light on the source of the discount. We focus on two subsamples of the full panel of firms: ongoing conglomerates and newly diversified firms. Table IV reports the results of estimating (18) for the subsample of firms that do not change the number of segments in which they operate. Specifically, we consider only the set of firms for which $s_t = s_{t-1} = \dots = s_{t-4}$, thus excluding all newly diversified (as well as refocused) firms from the sample. As Table IV documents, however, excluding these newly diversified firms does not eliminate the observed discount both in the data and in the model. Moreover, the actual value of the discount in our model is again very close to that observed by Lang and Stulz.

By contrast, Table V looks at the behavior of firms that change the numbers of segments of activity across adjacent years. Specifically, these firms are classified as "diversifying," if they change the number of sectors they operate in from one to two (formally $s_{t-1} = 1$ or 2 and $s_t = 3$) and as "focusing" firms if they reduce the number of activities from two to one ($s_{t-1} = 3$ and $s_t = 1$ or 2). These firms are then compared with those firms that keep the number of activities constant during the same period. For "diversifying" firms, the comparison group is the set of other previously focused firms that chose not to become

Table IV
Firms with Constant Segments

This table reports the results estimating the regression

$$q_{it} = b_0 + b_1 DIV_{it} + b_2 \ln(k_{it}) + \xi_{it},$$

on our artificial panel of firms. Here q_{it} is the value of Tobin's q for firm i at the beginning of period t , k_{it} is the beginning of period size of the firm, and DIV_{it} is a dummy variable that takes the value one if the firm is diversified in period t and is zero otherwise. The regression is performed only on the subsample of firms that do not change the number of segments in which they operate for a number of years. Specifically, we consider only firms for which $s_t = s_{t-1} = \dots = s_{t-4}$. In all cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics. The results of this estimation are then compared with the empirical findings from Table 8 in Lang and Stulz (1994).

Variable	Data	Model
<i>DIV</i>	-0.20	-0.17
(<i>t</i> -stat)	(-2.05)	(-3.14)
$\ln(k)$	-0.03	-0.66
(<i>t</i> -stat)	(-0.64)	(-3.48)

diversified in the current period. Similarly, focusing firms are compared with other diversified firms that chose to remain diversified.

Following Lang and Stulz (1994), we report two alternative results. First, we look at the average differences in q at the time that the firms choose to expand (or contract). Next, we also look at the dynamic effects of the decision, by comparing the effects of diversification (refocusing) on Δq .

The findings are somewhat inconclusive, both in the model and in the data, since no coefficient seems statistically significant.¹⁶ Nevertheless there is some suggestion, both in the model and in the data, that diversifying firms seem to experience drops in q (while the opposite happens for focusing firms). The model's implications for the level of q are somewhat less successful, but not statistically significant.

In recent empirical work, Burch and Nanda (2003) and Dittmar and Shivdasani (2003) find that firm spin-offs and divestitures tend to raise value, while improving the quality of investment. The results in Table V show that their evidence is also consistent with an optimal allocation of resources by firms. Intuitively, in our model, this occurs because firms will refocus when productivity shocks become more asymmetric. As a result, divestitures are associated, on average, with increases in productivity and investment in ongoing activities.

Overall, Tables IV and V broadly confirm the empirical success of our model. It is not only capable of generating a diversification discount, but also provides

¹⁶ Intuitively, the lack of statistical significance in the artificial sample is a consequence of the small variation in the shocks z . Although large variations would change this quantitative finding, a larger dispersion in z would lead to an unrealistically high dispersion in q .

Table V
Firms Changing Segments

This table compares firms that change the numbers of segments of activity across adjacent years with those firms that keep the number of activities constant. Specifically, firms are classified as “diversifying” if they change the number of sectors they operate from one to two (formally $s_{t-1} = 1$ or 2 and $s_t = 3$). We provide two separate results. First, we look at the average differences in q at the time the diversification takes place by estimating the regression

$$q_{it} = b_0 + b_1 DIV_{it} + \xi_{it},$$

for the subset of previously focused firms ($s_{t-1} < 3$). Here q_{it} is the value of Tobin’s q for firm i at the beginning of period t , and DIV_{it} is a dummy variable that takes the value one if firm has been focused at $t - 1$ and becomes diversified in period t and is zero otherwise. Next, we look at the dynamic effects of diversification, by comparing the effects of diversification on Δq . We accomplish that by estimating the following regression

$$\Delta q_{it} = b_0 + b_1 DIV_{it} + \xi_{it},$$

again only for the subset of previously focused firms. Here $\Delta q_{it} = q_{it} - q_{it-1}$. This table also reports the effects of refocusing on firm value, by estimating the same regressions as above, and letting DIV_{it} equal one if the firm has been diversified at $t - 1$ and becomes focused in period t and is zero otherwise. These regressions are only estimated for the subset of previously diversified firms. In all cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics. The results of this estimation are then compared with the empirical findings from Table 8 in Lang and Stulz (1994).

Variable	Regression on q_t		Regression on $\Delta q_t = q_{t+1} - q_t$	
	Data	Model	Data	Model
Diversifying Firms				
<i>DIV</i>	−0.163	0.045	−0.204	−0.038
(<i>t</i> -stat)	(−1.23)	(0.37)	(−1.60)	(−1.46)
Focusing Firms				
<i>DIV</i>	−0.016	0.035	0.024	0.020
(<i>t</i> -stat)	(−0.70)	(1.40)	(1.39)	(1.22)

quantitatively realistic results for the subsets of existing and newly diversified firms.

B.2. Robustness

While our benchmark calibration appears quite successful, it is interesting to examine the robustness of our findings to alternative choices of parameter values, particularly in light of the fact that there is relatively little a priori evidence for the four parameters, f , λ , σ , and ρ . Table VI investigates whether our main findings regarding the existence of a diversification discount are sensitive to our choices for these parameters. Specifically, Table VI compares the results of fitting the regression equation (18) to artificial samples, generated by varying our choices for the key parameters f , λ , σ , and ρ .

While the exact magnitude of the discount varies across the different experiments, the basic qualitative finding of a diversification discount seems to

Table VI
Robustness of the Diversification Discount

This table examines the robustness of our findings by reporting the results of estimating the following regression:

$$q_{it} = b_0 + b_1 DIV_{it} + b_2 \ln(k_{it}) + \xi_{it},$$

on several artificial panels of firms, obtained by varying the choice values for key parameters of the model. Here q_{it} is the value of Tobin's q for firm i at the beginning of period t , k_{it} is the beginning of period size of the firm, and DIV_{it} is a dummy variable that takes the value one if the firm is diversified in period t and is zero otherwise. In all cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics.

Variable	Baseline	$f = 0.004$	$\lambda = 0.5$	$\sigma = 0.02$	$\rho = 0.75$
<i>DIV</i>	-0.20	-0.03	-0.12	-0.07	-0.05
(<i>t</i> -stat)	(-5.39)	(-1.81)	(-3.55)	(-5.98)	(-7.02)
$\log(k)$	-0.70	-0.23	-0.70	-0.21	-0.27
(<i>t</i> -stat)	(-5.26)	(-26.95)	(-14.9)	(-41.13)	(-9.51)

be robust. The alternative values are chosen to indicate which changes lead to lower discounts. Thus, low variability in productivity (low σ or ρ) reduces the cross-sectional variability in q and thus the discount. High fixed costs, f , increase the cost savings of conglomerates, λf , and lower the discount. Decreasing the cost savings λ also lowers the implied discount. The reason is that lower synergies make diversification less attractive. With a low λ most conglomerates are formed to take advantage of decreasing returns to scale, and this effect is entirely captured by the large coefficient of $\ln k$.

C. Diversification and Productivity

In our model, productivity differentials play a key role in determining firm behavior and the observed link between diversification and firm valuation. In this section, we investigate whether the implied movements in firm and sectoral productivity are also consistent with existing empirical evidence. In a recent study, Schoar (2002) carefully documents the productivity patterns in manufacturing using the Longitudinal Research Database (LRD). Specifically, she computes Total Factor Productivity (TFP) for each plant, j , in each firm, i , and every period, t , by estimating the residual, ε_{ijt} , in the following log-linear Cobb–Douglas production function

$$\ln(y_{ijt}) = a_{jt} + b_{jt} \ln(k_{ijt}) + c_{jt} \ln(l_{ijt}) + \varepsilon_{ijt}. \quad (21)$$

Given this measure of productivity, we can examine the relation between firm diversification and firm productivity. Schoar (2002) focuses on two measures. First, she seeks to capture *static* differences in *average* productivity across firms by estimating the following equation:

$$TFP_{ijt} = a_1 + b_1 \times SEG_{it} + \mu_{ijt}, \quad (22)$$

Table VII
Diversification and Productivity

This table compares our findings with the results in Tables II and IV from Schoar (2002). First, we capture *static* differences in *average* productivity across firms by estimating the equation:

$$TFP_{ijt} = a_1 + b_1 \times SEG_{it} + \mu_{ijt}.$$

where TFP_{ijt} and SEG_{it} denote, respectively, total factor productivity in segment j and the logarithm of the number of segments in which firm i operates in period t . Second, the *dynamic* effects of diversification on *future* productivity are summarized with the regression:

$$TFP_{ijt} = a_2 + b_2 \times AFTER_{it} + v_{ijt},$$

where $AFTER_{it}$ is defined as a dummy variable that equals one in the period after the firm diversifies and is equal to zero otherwise. Results are reported for the benchmark case and an alternative calibration that matches Schoar's finding of 10 percent discount. In both cases we report the means across 100 simulations, for both the coefficients and the corresponding t -statistics.

Variable	Data	Model	
		Baseline	10% Discount
<i>SEG</i>	0.034	-0.023	0.009
(<i>t</i> -stat)	(2.13)	(-22.38)	(18.58)
<i>AFTER</i>	-0.026	-0.013	-0.008
(<i>t</i> -stat)	(-6.50)	(-7.81)	(-2.27)

where SEG_{it} is the logarithm of the number of segments in which firm i operates in period t . Thus, estimating $b_1 > 0$ implies that diversified (multisegment) firms are, on average, more productive than focused firms. In addition, she also examines the *dynamic* effects of diversification on *future* productivity. This is accomplished by estimating the equation

$$TFP_{ijt} = a_2 + b_2 \times AFTER_{it} + v_{ijt}, \quad (23)$$

where $AFTER_{it}$ is defined as a dummy variable that equals one in the period after the firm diversifies and it is equal to zero otherwise.¹⁷ Thus, a finding of $b_2 > 0$ implies that diversification *improves* plant productivity.

It is again relatively straightforward to use the artificial panel of firms generated by our model to replicate Schoar's (2002) procedures and compare the results. Given our measures of capital, labor, and output and assuming that each activity corresponds to one plant, we can easily estimate specifications (21)–(23). Table VII compares our findings with the results in Tables II and IV from Schoar.

While Schoar (2002) finds a significant productivity *premium* of more than 3 percent for diversified firms, our model implies that focused firms are, on average, 2.3 percent more productive. However, this result depends on the magnitude of the diversification discount, since lower productivity leads to lower

¹⁷ Schoar (2002) also adds variables such as age and the number of segments the firm operates. In the context of our model, however, age is not defined and the number of segments is redundant.

valuations. This is important since in Schoar's LRD sample the average market discount for diversified firms is only about 10 percent, while our model, which is calibrated to replicate the Lang and Stulz's (1994) results, implies a discount of about 20 percent. The last column of Table VII addresses this issue by recalibrating our model to generate a discount of exactly 10 percent, thus making our results directly comparable with hers. We find that in this case our model can also match the observed productivity *premium* for conglomerates.

Table VII also shows that our model successfully reproduces the observed *losses* of productivity *after* the firm diversifies.¹⁸ As Schoar (2002) argues, these findings reinforce the importance of distinguishing between the static effect of *being* diversified and the dynamic effect of *becoming* diversified. From a static or cross-sectional point of view, diversified firms are, on average, more productive than focused firms. However, as Figure 1 illustrates, diversification in our model is often the result of bad productivity shocks in ongoing activities. Thus, it is not surprising to find that, on average, diversification is associated with productivity losses in incumbent sectors, just as Schoar finds.¹⁹

These results suggest that our basic argument that diversification decisions are driven by efficient responses to productivity differentials is not the result of assuming unrealistic patterns for productivity. The fact that our model is consistent with much of the evidence also suggests a possible alternative interpretation to the more popular "new toy effect," that emphasizes a shift in focus by managers towards the newly acquired segments at the expense of incumbent ones. Our findings show that this evidence can also be rationalized in the context of a value maximizing model.

III. 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 firm diversification and performance. In our model, 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 synergies.

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.

¹⁸ Here, our results can only be compared with Schoar's (2002) estimates for incumbent plants since in our model new plants have no prior history.

¹⁹ Firms may also diversify if the diversification threshold moves because outside opportunities improve. In this case productivity in the incumbent sector need not fall for these firms.

We obtain several important results. First, we can show that firms currently expanding are not only less productive than other (nonexpanding) focused firms, but they also experience productivity losses after the expansion, as documented by Schoar (2002). Second, as Santalo (2001) and Maksimovic and Phillips (2002) suggest, we find that size differences can account for part of the differences both in productivity and in valuation across focused and diversifying firms. However, we also show that this size “effect,” cannot 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 as documented by Lang and Stulz (1994).

Appendix A.

A. Proofs

We derive all formal proofs under the most general set of conditions. Accordingly, define the set $S = \{1, 2, 3\}$. Let $K \times L \subseteq R_+^2$ be the space of inputs and suppose that the stochastic process for the shock has a bounded support $Z = [\underline{z}, \bar{z}] \times [\underline{z}, \bar{z}]$, $-\infty < \underline{z} < \bar{z} < \infty$. Moreover, define $\mathfrak{Z}_z$ and $\mathfrak{Z}_k$ as the minimal sigma-fields generated by Z and K , respectively. Finally let $F(k, l)$ denote a general decreasing returns to scale technology.

We make the following minimal assumptions regarding the nature of these functions and the size of the fixed costs.

ASSUMPTION 1: *The production function $F(\cdot)$: (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 $N(z_{t+1} | z_t) : Z \times \mathfrak{Z}_z \rightarrow [0, 1] \times [0, 1]$ that: (i) is stationary; (ii) is monotone; and (iii) satisfies the Feller property. Let $G(z)$ denote the invariant distribution of z .*

ASSUMPTION 3: *The fixed costs of production, f , are not too large, i.e., $\exists k \in R_+ : f \leq zF(k, l)$.*

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 $N(dz' | z)$ has the Feller property, it follows from Lemma 9.5 in Stokey and Lucas (1989) that

$$\int v(s', k', z') N(dz' | z) \in C(S \times K \times Z). \quad (\text{A1})$$

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

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

$$\int v_1(s', k', z') N(dz' | z) \geq \int v_2(s', k', z') N(dz' | z), \quad (\text{A2})$$

and hence

$$(Tv_1)(s, k, z) \geq (Tv_2)(s, k, z). \quad (\text{A3})$$

(b) **Discounting.**

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

$$(Tv + a)(s, k, z) = v(s, k, z) + \beta a = (Tv)(s, k, z) + \beta a. \quad (\text{A4})$$

Q.E.D.

A.1. Proof of Proposition 1

To show existence and uniqueness define the operator

$$(Tv)(s, k, z) = \max_{\{k', s'\}} \left\{ \pi(s', k, z) + (1 - \delta)k - k' + \beta \int v(s', k', z') N(dz' | z) \right\}, \quad (\text{A5})$$

$$s' \in \begin{cases} \{s, 3\}, & s = 1, 2 \\ S, & s = 3. \end{cases}$$

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:

- (a) $T : C(S \times K \times Z) \longrightarrow C(S \times K \times Z)$ (Lemma 1) and
- (b) 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 (A5).

Monotonicity then follows immediately from Theorems 9.7 and 9.11 in Stokey and Lucas (1989). Q.E.D.

A.2. Proof of Proposition 2

LEMMA 3: Let $\Pi(s, k, z)$ be defined by (A25). Then $\Pi(s, k, z)$ is (i) nonnegative; (ii) weakly increasing in k ; and (iii) decreasing in z^s and increasing in $z^{\tilde{s}}$, $s \neq \tilde{s}$.

Proof: (i) $\Pi(s, k, z) \geq 0$. By definition

$$\pi(3, k, z) = \pi(1, \theta^*k, z) + \pi(2, (1 - \theta^*)k, z) + \lambda f, \quad (\text{A6})$$

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

$$\pi(3, k, z) \equiv \begin{cases} \pi(1, k, z) - (1 - \lambda)f, & \theta^* = 1 \\ \pi(2, k, z) - (1 - \lambda)f, & \theta^* = 0, \end{cases} \quad (\text{A7})$$

since θ^* is chosen optimally, it follows that

$$\Pi(s, k, z) = \pi(3, k, z) - \pi(s, k, z) \geq 0. \quad (\text{A8})$$

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

$$\begin{aligned} \frac{\partial \pi(3, k, z)}{\partial k} &= \frac{\partial \pi(1, \theta^*k, z)}{\partial(\theta k)} \left(k \frac{\partial \theta^*}{\partial k} + \theta^* \right) \\ &\quad + \frac{\partial \pi(2, (1 - \theta^*)k, z)}{\partial(\theta k)} \left(-k \frac{\partial \theta^*}{\partial k} + (1 - \theta^*) \right). \end{aligned} \quad (\text{A9})$$

Noting that the optimal choice of θ^* implies

$$\frac{\partial \pi(1, \theta^*k, z)}{\partial(\theta k)} = \frac{\partial \pi(2, (1 - \theta^*)k, z)}{\partial(\theta k)}, \quad (\text{A10})$$

we immediately obtain

$$\frac{\partial \pi(3, k, z)}{\partial k} = \frac{\partial \pi(s', \theta^*k, z)}{\partial(\theta k)} \geq \frac{\partial \pi(s, k, z)}{\partial k}, \quad s = 1, 2, \quad (\text{A11})$$

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

$$\frac{\partial \Pi(s, k, z)}{\partial k} = \frac{\partial \pi(3, k, z)}{\partial k} - \frac{\partial \pi(s, k, z)}{\partial k} \geq 0, \quad s = 1, 2. \quad (\text{A12})$$

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

$$\frac{\partial \pi(3, k, z)}{\partial z^s} = \frac{\partial \pi(s, \theta^*k, z)}{\partial z^s} + \frac{\partial \pi(\tilde{s}, (1 - \theta^*)k, z)}{\partial z^s} = \frac{\partial \pi(s, \theta^*k, z)}{\partial z^s}, \quad (\text{A13})$$

since production in sector $\tilde{s}$ does not depend on the shock to sector i . Now, using the envelope theorem and the profits definitions (6) and (5) yields

$$\frac{\partial \pi(3, k, z)}{\partial z^s} = F(\theta^* k^s, \cdot) \leq F(k^s, \cdot) = \frac{\partial \pi(s, k, z)}{\partial z^s}, \quad (\text{A14})$$

and hence that

$$\frac{\partial \Pi(s, k, z)}{\partial z^s} \leq 0, \quad s = 1, 2. \quad (\text{A15})$$

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

LEMMA 4: *Let $\Psi(s, z)$ be defined by (A26). Then $\Psi(s, z)$ is: (i) nonnegative and (ii) decreasing in z^s and increasing in $z^{\tilde{s}}$, $s \neq \tilde{s}$.*

Proof: (i) $\Psi(s, z) \geq 0$. First note that

$$\begin{aligned} v(3, k', z') &= \max\{p(1, k', z'), p(2, k', z'), p(3, k', z')\} \\ &\geq \max\{p(s'', k', z'), p(3, k', z')\} = v(s', k', z'), \\ \forall (k', z') &\in K \times Z, \forall s'' \in \{1, 2\}. \end{aligned} \quad (\text{A16})$$

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

$$\begin{aligned} &\int \max\{p(1, k', z'), p(2, k', z'), p(3, k', z')\} N(dz' | z) \\ &\geq \int \max\{p(s'', k', z'), p(3, k', z')\} N(dz' | z). \end{aligned} \quad (\text{A17})$$

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

$$\beta \int v(3, k', z') N(dz' | z) - k' \geq \beta \int v(s', k', z') N(dz' | z) - k'. \quad (\text{A18})$$

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 $z^{\tilde{s}}$, $s \neq \tilde{s}$. Suppose $z^s \gg z^{\tilde{s}}$. Then

$$p(s, k, z) > p(\tilde{s}, k, z), \quad (\text{A19})$$

and consequently

$$v(3, k, z) \approx \max\{p(s, k, z), p(3, k, z)\}. \quad (\text{A20})$$

Given the monotonicity of $N(\cdot)$, it follows that

$$\begin{aligned} \int v(3, k', z') N(dz' | z) &\approx \int \max\{p(s, k', z'), p(3, k', z')\} N(dz' | z) \\ &= \int v(s, k', z') N(dz' | z), \end{aligned} \quad (\text{A21})$$

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

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

$$v(3, k, z) \approx \max\{p(\bar{s}, k, z), p(3, k, z)\} \quad (\text{A22})$$

and

$$\begin{aligned} \int v(3, k', z') N(dz' | z) &\approx \int \max\{p(\bar{s}, k', z'), p(3, k', z')\} N(dz' | z) \\ &> \int v(s, k', z') N(dz' | z), \end{aligned} \quad (\text{A23})$$

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

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

Using equation (9) we can rewrite the optimal diversification decision of a focused firm as

$$\Pi(s, k, z) + \Psi(s, z) \geq (1 - \lambda)f, \quad s \in \{1, 2\}, \quad (\text{A24})$$

where

$$\Pi(s, k, z) \equiv \pi(3, k, z) - \pi(s, k, z) + (1 - \lambda)f, \quad (\text{A25})$$

and

$$\begin{aligned} \Psi(s, z) = \Psi(s', z)|_{s'=s} &\equiv \max_{k'} \left\{ \beta \int v(3, k', z') N(dz' | z) - k' \right\} \\ &\quad - \max_{k'} \left\{ \beta \int v(s, k', z') N(dz' | z) - k' \right\}. \end{aligned} \quad (\text{A26})$$

Equation (A24) decomposes the optimal diversification decision into a “profit” component, $\Pi(s, k, z)$, and an “option” component, $\Psi(s, z)$, associated with the continuation payoffs. By definition $\hat{k}(s, z)$, satisfies

$$\Pi(s, \hat{k}(s, z), z) + \Psi(s, z) = (1 - \lambda)f, \quad \forall (s, z) \in S \times Z. \quad (\text{A27})$$

Lemmas 3 and 4 imply that the left-hand side is nonnegative 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 $\hat{k}(s, z) = 0$.

To establish monotonicity let $z = (z^s, z^{\bar{s}})$ and $\hat{z} = (z^s + \Delta z^s, z^{\bar{s}})$, with $\Delta z^s > 0$. It follows from (A27) that

$$\Pi(s, \hat{k}(s, \hat{z}), \hat{z}) + \Psi(s, \hat{z}) = (1 - \lambda)f. \quad (\text{A28})$$

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

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

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

$$p(3, \hat{k}(s, z), z) \geq \max\{p(1, \hat{k}(s, z), z), p(2, \hat{k}(s, z), z)\}, \quad (\text{A29})$$

or, simply, by

$$\min_{s \in \{1, 2\}} \{\Pi(s, \hat{k}(s, z), z) + \Psi(s, z)\} = (1 - \lambda)f, \quad \forall z \in Z. \quad (\text{A30})$$

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

Proof of Corollary 1: In the absence of fixed costs inequality (A24) is always satisfied. Q.E.D.

Proof of Corollary 2: Inequality (A24) is always satisfied if $\lambda \geq 1$. Q.E.D.

Appendix B

A. Solution Method

The computational strategy involves the following steps:

1. Solving the Bellman equation (8) and computing the optimal firm decision rules.
2. Using the optimal decision rules to iterate on (13) and compute the stationary measure $\mu = \mu' = \mu^*$.
3. Computing aggregate quantities and using the market clearing condition (16) to determine the equilibrium levels of consumption and labor.

Given the properties of our problem, the first step is better implemented with the less efficient but more robust method of value function iteration on

a discrete state space. We specify a grid with a finite number of points for the capital stock as well as a finite approximation to the normal random vector z . The latter task is accomplished using Tauchen and Hussey's (1991) method for optimal discrete state space approximations to normal random variables. We use 15×15 grid points for this procedure. The space for the capital stock is divided into 201 equally spaced elements. In either case the results were relatively unchanged when we used finer grids. The upper bound for capacity, k , was chosen to be nonbinding at all times.

To compute μ^* , we take the optimal value function $v(s, k, z)$ and the decision rules $\mathbf{k}(s, k, z)$ and $\mathbf{s}(s, k, z)$, as well as the stochastic process for the technology shocks z and proceed as follows:

- Define the size of the panel data, by specifying the number of firms M and the length of time T .
- Simulate a sequence of exogenous technology shocks $z_{it} = (z_{it}^1, z_{it}^2)$ for each firm i in every period t .
- For the initial period
 - (i) Initiate each firm's capital stock at $k = k_0$.
 - (ii) Start the simulation by using draws from a uniform distribution to randomly allocate firms to either sector one or two.
- For all other periods
 - (i) Given the current state for each firm i , $(s_{it-1}, k_{it}, z_{it})$ use the optimal policy functions to determine the next period's capital stock, k_{it+1} , and sectoral decision, s_{it} .
 - (ii) Using the value function, compute the current market value of the firm i , v_{it} .
 - (iii) Using the stochastic process for z , compute next period's shock z_{it+1} .
 - (iv) Construct the cross-sectional distribution of firms $\mu_{it} = \mu(s_{it}, k_{it}, z_{it})$.
- Continue the simulation until $||\mu_{it} - \mu_{it+1}|| < \varepsilon$.

Using the stationary distribution, μ , it is straightforward to use the goods market condition to obtain aggregate consumption.

REFERENCES

- Amihud, Yakov, and Baruch, Lev, 1981, Risk reduction as a managerial motive for conglomerate mergers, *Bell Journal of Economics* 12, 605–617.
- Berger, Philip, and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–65.
- Bernardo, Antonio, and Bhagwan Chowdhry, 2002, Resources, real options and corporate strategy, *Journal of Financial Economics* 63, 211–234.
- Burch, Timothy, and Vikram Nanda, 2003, Divisional diversity and the conglomerate discount: The evidence from spin-offs, *Journal of Financial Economics* 70, 69–98.
- Burnside, Craig, 1996, Production function regressions, returns to scale and externalities, *Journal of Monetary Economics* 37, 177–201.
- Campa, Jose, and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 135–160.

- Chevalier, Judith, 2001, Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms, Unpublished manuscript, University of Chicago.
- Denis, Dabid, Dianne Denis, and Atulya Sarin, 1997, Agency problems, equity ownership and corporate diversification, *Journal of Finance* 52, 135–160.
- Dittmar, Amy, and Anil Shivdasani, 2003, Divestitures and divisional investment policies, *Journal of Finance* 58, 2711–2743.
- Gomes, Joao, 2001, Financing investment, *American Economic Review* 91, 1263–1285.
- Graham, John, Michael Lemmon, and Jack Wolf, 2002, Does corporate diversification destroy value? *Journal of Finance* 57, 695–720.
- Hopenhayn, Hugo, 1992, Entry, exit, and firm dynamics in long run equilibrium, *Econometrica* 60, 1127–1150.
- Inderst, Roman, and Holger Muller, 2003, Internal versus external financing: an optimal contracting approach, *Journal of Finance* 58, 1033–1062.
- Jensen, Michael, 1986, Agency costs of free cash flow, corporate finance, and takeovers, *American Economic Review* 76, 323–329.
- Jensen, Michael, and Kevin Murphy, 1990, Performance pay and top management incentives, *Journal of Political Economy* 98, 225–264.
- Lamont, Owen, and Christopher Polk, 2001, The diversification discount: Cash flows versus returns, *Journal of Finance* 56, 1693–1721.
- Lamont, Owen, and Christopher Polk, 2002, Does diversification destroy value? Evidence from industry shocks, *Journal of Financial Economics* 63, 51–77.
- Lang, Larry, and Rene Stulz, 1994, Tobin's q , corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248–1280.
- Maksimovic, Vojislav, and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently across industries? *Journal of Finance* 57, 721–767.
- Matsusaka, John, 2001, Corporate diversification, value maximization, and organizational capabilities, *Journal of Business* 74, 409–431.
- Rajan, Raghuram, Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance* 55, 35–80.
- Santalo, Juan, 2001, Diversification discount, size discount and organizational capital, Unpublished manuscript, University of Chicago.
- Scharfstein, David, and Jeremy Stein, 2000, The dark side of internal capital markets I, *Journal of Finance* 55, 2537–2564.
- Schleifer, Andrei, and Robert Vishny, 1989, Managerial entrenchment, the case of manager-specific investment, *Journal of Financial Economics* 25, 123–139.
- Schoar, Antoinette, 2002, Effects of corporate diversification on productivity, *Journal of Finance* 57, 2379–2403.
- Stokey, Nancy, and Robert Lucas, 1989, *Recursive Methods in Economic Dynamics* (Harvard University Press, Cambridge, MA).
- Stulz, Rene, 1990, Managerial discretion and optimal financing policies, *Journal of Financial Economics* 26, 3–27.
- Villalonga, Belen, 2001, Does diversification cause the “diversification discount”? Unpublished manuscript, Harvard University.
- Wernerfelt, Birger, and Cynthia A. Montgomery, 1988, Tobin's q and the importance of focus in firm performance, *American Economic Review* 78, 246–250.
- Whited, Toni, 2001, Is it inefficient investment that causes the diversification discount? *Journal of Finance* 56, 1667–1691.