

# Are Conglomerates Under-Diversified?\*

Fernando Anjos<sup>†</sup>

Cesare Fracassi<sup>‡</sup>

## Abstract

Mergers are an important economic activity, but informational frictions may limit the extent to which this form of resource reallocation takes place, especially when combining firms across less related—or distant—industries. This paper conjectures that the presence of well-diversified conglomerates in the economy generates an externality that facilitates the occurrence of value-adding distant mergers. In this context we study how efficient the equilibrium level of conglomerate diversification is, by developing a dynamic industry-spatial model of merger activity. According to the model, conglomerates can be either under- or over-diversified in equilibrium, depending on parameter values. We calibrate the model using data on conglomerate valuations and an input-output-based measure of industry distance. The calibrated model yields a good fit to data and implies that diversified firms operate on average with a low segment distance compared to the optimum.

First draft: November 18, 2011

This version: November 14, 2012

**JEL classification:** D2, D57, G34, L14, L25.

**Keywords:** corporate diversification, mergers, matching, input-output tables, networks.

---

\*The authors thank comments from and discussions with Andres Almazan, Aydoğan Altı, Jonathan Cohn, Richard Lowery, Ramesh Rao, Ray Reagans, Alessio Saretto, Laura Starks, and Malcolm Wardlaw. The authors also thank comments from seminar participants at the University of Texas at Austin and conference participants at the 2012 European meetings of the Econometric Society (under the title “Optimal Segment Distance in Diversified Firms”).

<sup>†</sup>University of Texas at Austin, McCombs School of Business, 1 University Station B6600, Austin TX 78712. Telephone: (512) 232-6825. E-mail: fernando.anjos@mcombs.utexas.edu

<sup>‡</sup>University of Texas at Austin, McCombs School of Business, 1 University Station B6600, Austin TX 78712. Telephone: (512) 232-6843. E-mail: cesare.fracassi@mcombs.utexas.edu

# Are Conglomerates Under-Diversified?

## Abstract

Mergers are an important economic activity, but informational frictions may limit the extent to which this form of resource reallocation takes place, especially when combining firms across less related—or distant—industries. This paper conjectures that the presence of well-diversified conglomerates in the economy generates an externality that facilitates the occurrence of value-adding distant mergers. In this context we study how efficient the equilibrium level of conglomerate diversification is, by developing a dynamic industry-spatial model of merger activity. According to the model, conglomerates can be either under- or over-diversified in equilibrium, depending on parameter values. We calibrate the model using data on conglomerate valuations and an input-output-based measure of industry distance. The calibrated model yields a good fit to data and implies that diversified firms operate on average with a low segment distance compared to the optimum.

**JEL classification:** D2, D57, L14, L25, G34.

**Keywords:** corporate diversification, mergers, matching, input-output tables, networks.

# 1 Introduction

Merger activity represents a key mechanism through which the economy reallocates resources, and the importance of this phenomenon has prompted much research on the topic, both from a micro- and macro-economic perspective. Merger deals may take many forms, but irrespective of the details, this is a setting where uncertainty about payoffs and concerns about adverse selection are pervasive, and where in all likelihood these informational issues prevent a number of potentially efficient deals from taking place. Informational problems would be especially relevant for deals involving firms from less related—or distant—industries, where uncertainty about synergies and asymmetric information are more severe.

Our paper argues that the presence of informational frictions stems partly from the economic environment itself. We specifically conjecture that the presence of well-diversified conglomerates in the economy generates an externality that potentially alleviates these frictions, and whereas we do not isolate the mechanisms through which this externality occurs, we believe various could be at work. A few examples illustrate our main idea. First, consider an example without asymmetric information, but where agents are highly uncertain about the synergy gains associated with a yet-untried combination of businesses from two particular sectors. In this setting, the first pair of firms to complete such a deal—becoming in the process a diversified corporation—generates an informational spillover (however imperfect) about synergy gains, and this knowledge externality can actually trigger other, similar, deals. Second, consider a setting where there is no uncertainty about synergies, but where there is asymmetric information about the value of firms’ industry-specific assets-in-place. In particular let us consider the role of manager specialization. The more specialized managers become, the lower ability they have to evaluate assets from other industries, which due to adverse selection may preclude a deal from taking place. This effect would be partially offset in an economy with many diversified corporations, as long as managers with past experience working in conglomerates are inherently less specialized, and therefore relatively more skilled

in valuing assets from different sectors. Finally, it could also be the case that the social connections formed inside a diversified firm act as an informational shortcut in the economy, making it easier to transmit reliable signals about asset quality across distant sectors via inter-personal connections.

Our main research question is much in the spirit of macro-economic literature on labor markets with search frictions (Diamond, 1993; Mortensen and Pissarides, 1994). While the latter focuses on the question of whether equilibrium unemployment is too high or too low—relative to the socially desirable outcome—, we ask whether conglomerates in the economy are under- or over-diversified. We address our research question by building a dynamic industry-spatial model of corporate diversification, which we calibrate using data on conglomerate valuations and the input-output-based measure of industry distance from Anjos and Fracassi (2011). Our model provides a good fit to data and suggests that in equilibrium conglomerates may be significantly under-diversified. This result notwithstanding, the calibration also points to the inefficiency not being quantitatively significant.

Our modeling approach has three distinctive features: (i) Diversification adds value by potentially enhancing resource reallocation in a way that is hard for specialized firms and external markets to achieve, but at the cost of a potentially entrenched (and costly) top management; (ii) Consistent with empirical evidence,<sup>1</sup> diversification occurs via merger events, which are modeled using a search-and-matching framework along the lines of Diamond (1993) and Mortensen and Pissarides (1994); (iii) The pervasiveness of well-diversified conglomerates influences the (merger) matching process that generates diversification itself. In particular, we assume that conglomerates operating in segments that are relatively distant—in an industry-spatial sense—makes it more likely that specialized firms distant from one another are matched. The result that under-diversification occurs in equilibrium can then be understood through the lens of standard coordination-failure arguments: If (a) there are too

---

<sup>1</sup>In Graham, Lemmon, and Wolf (2002) almost two thirds of the firms that increase the number of segments implement this strategy via acquisition.

few conglomerates operating at high distance, then (b) specialized firms operating at high distance rarely match, which in turn leads to (a).

The above intuition notwithstanding, the model also predicts that conglomerates may be over-diversified in equilibrium. This distortion occurs when the baseline matching technology—that is, absent the externality channel—already produces many matches at a distance that is greater than the *ideal* distance, from the perspective of maximizing diversification synergies. In this case, adding the externality further dislocates average matches to a region of industry distance that is not advantageous, and thus the externality has negative effects. Our calibration exercise allows us to learn about the relationship between the synergy technology and the matching technology, and it is this relationship that according to the model ultimately determines whether conglomerates are under- or over-diversified.

From an empirical standpoint we start by recalling the finding in Anjos and Fracassi (2011) that segment distance is on average positively associated with conglomerate value, which in light of previous literature is perhaps surprising. In a multivariate regression setting this pattern is robust to including standard variables used in diversification studies, as well as imposing firm fixed effects. We further document that most diversified firms operate at relatively low distances, which begs the question of why firms are apparently not (fully) taking advantage of the segment-distance effect. Our model rationalizes this pattern, and in particular it shows that this naturally obtains when the level of diversification is inefficiently low in equilibrium.

Our approach to diversification-associated synergies plays an important role in our results, and in particular in explaining the observed relationship between segment distance and conglomerate value. The key assumption here is that there is some uncertainty about the “skills”, or type of capital/resources, required for successfully implementing projects. Whereas there is nothing a specialized firm can do about this, a conglomerate has the option to internally reallocate resources in order to minimize this problem. Thus our approach is

close to the internal capital markets literature (Stein, 1997; Scharfstein and Stein, 2000), but we consider an ability to reallocate resources other than financial capital and which are industry-specific. The fundamental implication of our model for synergies is that these exhibit an inverted-U shape with respect to segment distance, and that is why when firms cluster around low distances one observes a positive association between distance and value. The intuition for the non-monotonic behavior is that for low distances firm skills/resources are too similar, so although many times it is the case that a reallocation adds value, such value is very low. On the other hand, if distances become too large, there are very few occurrences where one segment's skills/resources are actually helpful for projects occurring in other segments.

The result that conglomerates are under-diversified stands in stark contrast with the position of earlier literature on corporate diversification. Indeed, early finance literature on this topic conveyed the idea not only that there are too many conglomerates, but also that conglomerates have too many segments and participate in industries that are too unrelated (Lang and Stulz, 1994; Berger and Ofek, 1995). Presumably this would result from managers having a preference for earnings smoothing or being overconfident, and thus not acting in the shareholders' best interests. The notion of empire-building and unfocused managers seemed consistent with the empirical regularity of a large "diversification discount", plus the observation of additional "discounts" associated with number of segments and low relatedness. The normative prescription quickly became: specialize and refocus. But as we show in our paper, this conclusion may be erroneous once externalities are taken into account. It is also interesting to note that our under-diversification result coexists with the model delivering an apparent diversification discount,<sup>2</sup> although corporate diversification actually accounts for 5-20% of firm value in our calibration.<sup>3</sup>

---

<sup>2</sup>Other authors have proposed dynamic models where an apparent diversification discount exists, most notably Gomes and Livdan (2004). The mechanism that generates an apparent discount in our paper is very close to the one proposed in Anjos (2010).

<sup>3</sup>The lower bound of 5% corresponds to the value of the diversification option for a specialized firm; the

Whereas the existence of a diversification discount is no longer an established consensus (Villalonga, 2004; Custódio, 2009), and recent empirical research has even suggested novel bright sides for diversification (Anjos and Fracassi, 2011; Tate and Yang, 2011), the literature still has ignored the more macro/general-equilibrium aspects of diversification; which we view as the main contribution of our paper.

Our paper also relates to the strand of diversification research that attempts to understand what kind of segment-combination strategies work (e.g., related/unrelated). Berger and Ofek (1995) investigate whether segments belonging to the same SIC code affects firm value, Fan and Lang (2000) study the effects of vertical relatedness using input-output-based measures, and Hoberg and Phillips (2010) use text-based analysis of 10K product descriptions to measure similarity across firms. These papers all view relatedness as having positive economic effects. We start our paper by revisiting the concept of relatedness and suggesting the segment distance variable in Anjos and Fracassi (2011)—based on the overall architecture of input-output flows—as a proxy for the inverse of relatedness. The logic of other relatedness measures notwithstanding, we believe our measure is an appropriate proxy for the value-creation potential of a particular combination of organizational skills in a diversified business, and we argue this in more detail in the main body of the paper.

The paper proceeds as follows. Section 2 presents some initial evidence on segment distance and the value of diversified firms. Section 3 develops the theoretical setup, which entails a model for the relationship between segment distance and static/flow synergies; and a model for the process through which diversification activity occurs and firm boundaries change. Section 4 calibrates the dynamic model and presents the main results of the paper. Section 5 concludes. An appendix contains all proofs and summary statistics.

---

upper bound of 20% corresponds to a diversified firm with segments at the ideal distance, at the time the diversifying merger occurs.

## 2 Motivating data

### 2.1 The segment-distance variable

Our basic conceptual approach follows from standard economics, in that we view firms as possessing a collection of different types of productive capital (resources), which are combined in order to generate output. Our main argument builds on the idea that a similarity, or distance, measure across resources can be constructed, and that such distance is economically meaningful. Being more precise, we take the following threefold view (detailed in the model section): (i) organizational resources, or skills, are critical in terms of firm profitability and value; (ii) the right *combination* of organizational skills is important, and in particular resource distance within the firm matters; (iii) merger activity and corporate diversification are important forms of achieving the right combination of organizational skills. By organizational skills we mean some combination of managerial talent and the different forms of firm-specific organizational capital—arising from learning curves in a specific business context—that accumulates over time.

The form of capital we have in mind is summarized by a firm’s ability to efficiently operate in a certain business context, and, unlike the relatedness approach from other authors, it does not necessarily require that two sectors are technologically similar: for example, one can envision a management team of a particular company A being effective operating a division in sector B, if the two sectors are importantly affected by what happens to industry C; even if A and B are neither rivals nor have a direct supply-chain relationship.

The question then is: how to measure similarity across organizational resources? From an empirical standpoint, the challenge is to obtain a proxy that reflects the conceptual notion elaborated above, and, at the same time, does not pick up effects related to market-power considerations or issues related to incomplete contracting. An empirical measure that achieves these objectives is the segment-distance variable in Anjos and Fracassi (2011). This

paper takes a networks approach to constructing segment distance, i.e., it considers direct and indirect linkages in the input-output space in order to determine the pairwise industry distance for all possible pairs of industries in the economy. Using this network methodology, a small cross-industry distance may occur because there is a strong direct customer-supplier relationship, but also if two industries that are apparently disconnected share many of the same customer-industries and/or many of the same supplier-industries; one could say in the latter case that these two industries share the same business environment despite being bilaterally disconnected. With a networks approach to distance, one can thus distinguish the effect of proximity in organizational skills—a firm’s ability to operate efficiently in a particular business environment—from more-standard arguments for vertical integration, which we argue are more appropriately proxied for by the intensity of the direct connection.<sup>4</sup> The segment distance variable, for a conglomerate, is defined as follows:

$$Seg.Distance = \frac{\sum_{i \in \mathcal{I}} \sum_{j > i \wedge i \in \mathcal{I}} l_{ij}}{M(M-1)/2}, \quad (1)$$

where  $\mathcal{I}$  denotes the set of industries a diversified firm participates in,  $M$  is the size of this set, and  $l_{ij}$  the length of the shortest path between industries  $i$  and  $j$ . This shortest path is computed by considering the overall industry network of the economy, and this is why indirect linkages matter. The reader is referred to Anjos and Fracassi (2011) for further details.

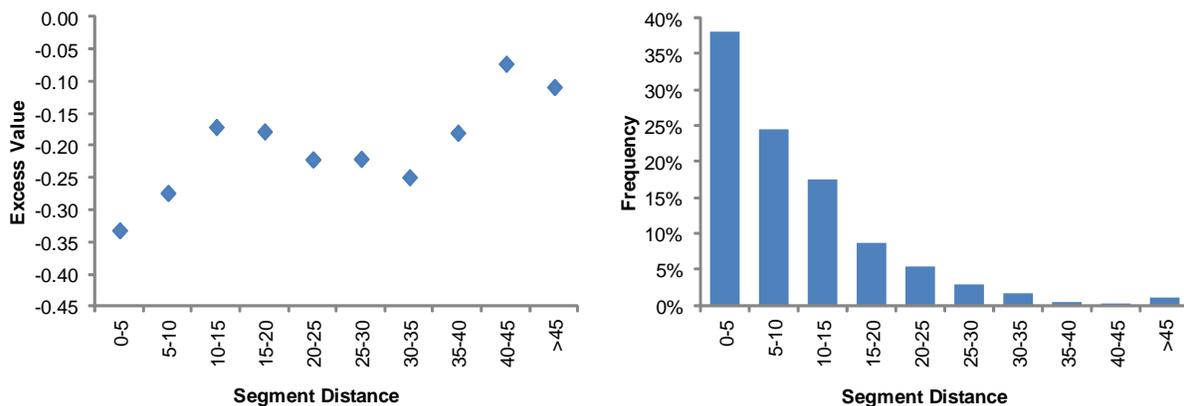
## 2.2 Segment distance and conglomerate value

One of the distinctive features of the model we develop is that segment distance determines the size of diversification-associated synergies. If this assumption holds true in data, then one should in principle observe a relationship between segment distance and firm value.

---

<sup>4</sup>And indeed, Ahern and Harford (2012) show empirically that the intensity in bilateral input-output flows is an important determinant of cross-industry merger activity.

In this section we make a first pass in documenting this relationship, using data on U.S. firms from 1990 to 2008. Our key independent variable, segment distance, is computed using the detailed input-output tables for the year 1997, approximately the middle of the sample period. Our key dependent variable is conglomerate *excess value*, which we compute following Berger and Ofek (1995), as do many other papers on corporate diversification. This variable corresponds to the log-difference of the conglomerate’s Tobin’s  $Q$  and the Tobin’s  $Q$  of a similar portfolio of specialized firms; the idea behind this variable is to control for industry-specific valuation patterns. Summary statistics are presented in table A.5 in the appendix.



**Figure 1: Segment Distance and Excess Value.** The left panel shows conglomerate average excess value, conditional on segment-distance class. The right panel shows segment-distance distribution. Excess value is defined as the log-difference between the Tobin’s  $Q$  of a conglomerate and the Tobin’s  $Q$  of a similar portfolio of specialized firms, following Berger and Ofek (1995). Segment Distance is the average input-output-based distance across conglomerate segments, following Anjos and Fracassi (2011).

The left panel in figure 1 shows the empirical association between segment distance and conglomerate valuation, plotting in-sample average excess value for ten segment-distance intervals. It is perhaps puzzling that excess value is on average negative for every segment-distance class—the celebrated *diversification discount*—but actually our model accommodates this feature even though firms are rational value maximizers and diversification synergies are positive. The relationship between segment distance and firm value is economically significant; for example, going from the 20-th segment-distance decile to the 80-th decile cor-

responds to an increase of approximately 13% in average firm value, as measured by Tobin’s  $Q$ .

The right panel of figure 1 shows that most conglomerates display a relatively low segment distance, at least if one takes into account the apparent gains of operating at higher distances suggested by the relationship between segment distance and excess value. Admittedly, many endogeneity concerns could drive these results, but it is surprising to find conglomerates operating in “unrelated”—i.e., distant—businesses being the ones that on average draw the most synergies; this certainly stands in contrast with the early consensus on the effect of relatedness on firm value. However, our model shows that it is precisely when firms cluster at low distances that one should expect to observe a positive association between segment distance and excess value. Furthermore, this is also the instance where—according to our model—it is highly likely that segment distance is inefficiently low from a social welfare perspective.

The simplistic analysis in the left panel of figure 1 is naturally subject to many endogeneity concerns. Whereas we cannot address all of these, we can rule out some simple alternative explanations that would render the association spurious. Table 1 conducts a multivariate regression analysis, with excess value as the dependent variable, where we investigate how robust the segment distance effect is.

Specification (1) in table 1 presents the correlation between segment distance and excess value, but now controlling for year fixed effects; model (2) adds firm fixed effects. In both models a positive association between segment distance and excess value is still present. Specification (3) adds control variables that are common in the diversification literature: number of segments and number of related segments (the relatedness measure in Berger and Ofek (1995)). It also includes a vertical relatedness measure, computed following Fan and Lang (2000), which allows us to differentiate our story from more-standard arguments related to vertical integration. We note that the vertical relatedness loads only on the intensity of

**Table 1: Excess Value and Segment Distance.** The dependent variable is Excess Value, defined as the log-difference between the Tobin's  $Q$  of a conglomerate and the Tobin's  $Q$  of a similar portfolio of specialized firms, following Berger and Ofek (1995). The table presents ordinary least squares regression coefficients and robust t-statistics clustered at the conglomerate level. The main explanatory variable is Segment Distance, defined as the average level of binary distance for every possible pair of industries that the conglomerate participates in, using the 6-digit Input-Output industry classification system. All variables are defined in detail in the appendix. A constant is included in each specification but not reported in the table. Inclusion of fixed effects is indicated at the end. Significance at 10%, 5%, and 1%, is indicated by \*, \*\*, and \*\*\*.

	(1)	(2)	(3)	(4)	(5)	(6)
Segment Distance	0.004*** (3.39)	0.005* (1.85)	0.002* (1.71)	0.005* (1.81)	0.002* (1.67)	0.005** (1.99)
N. Segments			-0.015 (-1.07)	-0.095*** (-4.37)	0.002 (0.15)	-0.075*** (-3.48)
Related Segments			0.052** (2.00)	0.041 (0.97)	0.065** (2.47)	0.035 (0.81)
Vert. Relatedness			-0.002 (-0.99)	0.072 (1.20)	-0.001 (-0.76)	0.072 (1.19)
Excess Centrality			1.563** (2.26)	2.243*** (2.62)	1.618** (2.28)	1.935** (2.22)
Log Assets					-0.019** (-2.49)	-0.163*** (-7.23)
EBIT/Sales					-0.005** (-2.35)	-0.001 (-0.58)
CAPEX/Sales					0.003 (0.76)	0.052** (2.29)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes
$R^2$	0.014	0.020	0.018	0.031	0.024	0.053
Obs.	7,892	7,892	7,892	7,892	7,590	7,590

*direct* bilateral links. Finally, model (3) also includes the excess centrality measure in Anjos and Fracassi (2011), which captures a conglomerate's informational advantage relative to specialized firms. Once more, the coefficient of segment distance remains statistically and economically significant, and this is also true once firm fixed effects are added (model (4)). Specifications (5) and (6) add financial variables that are popular in regression approaches to corporate diversification, and the segment-distance effect still obtains.

The multivariate regression approach establishes that the association between segment

distance and excess value is not driven by other firm characteristics that have been documented to also affect firm value. It is perhaps worthwhile to emphasize that the segment-distance effect survives the inclusion of firm fixed effects. This allows us to rule out an explanation based on persistent managerial skill or unobserved organizational capital, where better firms are the ones that simultaneously are more profitable running their businesses and also have more ability to evaluate merger/expansion opportunities at a distance.<sup>5</sup>

It is interesting to note that according to table 1 vertical relatedness does not seem to generate higher firm valuations. This does not mean, however, that firms do not take advantage from vertical relatedness: it is perfectly consistent with firms competing away all potential gains associated with vertical integration.

### 3 Model setup

This section contains our theoretical setup. The setup is built in the spirit of many dynamic models in economics; there is an equilibrium model for flow payoffs which is embedded in a dynamic setup. We start by developing the static equilibrium model for flow payoffs (section 3.1), and afterwards we develop the search-and-matching approach to diversifying activity (section 3.2). The solution to the model is presented in section 3.3.

#### 3.1 Flow payoffs (static setup)

The economy comprises a continuum of business units (henceforth BUs), where BU  $i$  is characterized by a location  $\alpha_i$  on a circle with measure 1.<sup>6</sup> The different locations in the circle represent, in a simplified way, the different industries in the economy; or, more precisely, the type of skills/competences/capital that is mostly associated with each industry. We refer

---

<sup>5</sup>With the caveat that time-varying managerial skills or firm organizational capital could still render our results spurious.

<sup>6</sup>The advantage of working with a circle (instead of a line, for example) is that this makes the solution to the matching model very tractable, given the symmetry of the circle.

to this space of productive location as the *skill-type* circle.

Business units are organized either as a single-BU specialized firm or as a two-BU (or two-segments) corporation, which we term a conglomerate. We take the organizational forms as given for now; these are endogenized later (section 3.2). The next two subsections further characterize the static payoffs of specialized and diversified firms.

### 3.1.1 Specialized firms

Each BU in the economy is presented exactly with one project, and this project is also characterized by a location in the skill-type circle, denoted by  $\alpha_{P_i}$ . This location represents the *ideal* skill type, that is, the skill type that maximizes the project's profitability. The location of the project is drawn from a uniform distribution centered at  $\alpha_i$ —which implies on average BUs are well-equipped to implement the projects they are presented with—, with lower bound  $\alpha_i - \sigma$  and upper bound  $\alpha_i + \sigma$ . The higher  $\sigma$  is, the higher the uncertainty about the ideal skill type required by projects. For tractability we assume  $\sigma < 1/4$ , which greatly simplifies the analysis.<sup>7</sup> The support of the distribution for project location corresponds to the dashed arc in figure 2.

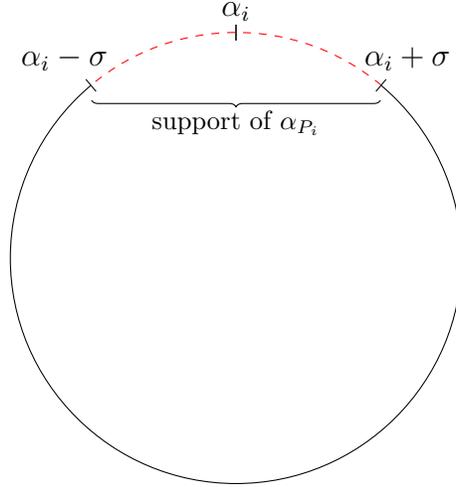
If BU  $i$  is organized as a specialized firm, then its profit function is given by the following expression:

$$\pi_i = 1 - \phi z_{i,P_i}, \tag{2}$$

where  $z_{i,P_i}$  is the length of the shortest arc connecting  $\alpha_i$  and  $\alpha_{P_i}$ , that is, the distance between the skill type of the BU and the ideal skill type required by the project. Parameter  $\phi > 0$  gauges the cost of project-firm mismatch. It follows then from our assumptions that

---

<sup>7</sup>Tractability with low enough uncertainty about project location originates from the fact that we only have to consider one-sided overlap in project-generating regions. The advantage of this assumption is clear in the derivations and proofs presented in the appendix. We also believe this assumption is fairly innocuous in terms of the main results.



**Figure 2:** Skill-type location of firms and projects.

the expected profits of a profit-maximizing single-BU firm, denoted as  $\pi_0$ , are given by

$$\pi_0 := \mathbb{E}[\pi_i] = 1 - \phi \frac{\sigma}{2}. \quad (3)$$

Finally we assume that projects cannot be traded across firms. This could be due, for example, to adverse selection; and would be consistent with interpreting the boundaries of the firm as information boundaries (as suggested, e.g., in Chou (2007)). The stark representation of these informational costs (zero inside firms and infinite across firms) is made in order to keep the model tractable; it should be interpreted as the costs being higher for information transmission across firms, relative to information transmission inside firms.

### 3.1.2 Diversified firms

To keep the framework tractable, the only form of corporate diversification we consider is a company comprising two segments. If BU  $i$  is part of the same firm as BU  $j$ , then profits are similar to those of a specialized firm with the exception that projects can be traded (swapped) inside the firm; and this ex-post choice is assumed to be made optimally by the *headquarters* (henceforth HQ) of the multi-segment firm. We assume the HQ of a multi-

segment firm imposes additional costs on the firm, but we postpone this description until section 3.2. This mechanism of internal project trade aims to represent the advantage of having access to an internal pool of resources that the firm can deploy in an efficient way, given the business environment the firm is facing (here, the “project”) and the nature of which is imperfectly known ex ante.

Next we derive the expected (gross) profit function for a diversified firm, taking segment distance as given. This is presented in proposition 1.

**Proposition 1** *The expected gross profit of a BU in a diversified firm with segments located at distance  $z$ , denoted by  $\pi_1^G(z)$ , is given by the following expressions:*

1. If  $z \leq \sigma$ ,

$$\pi_1^G(z) = 1 - \phi \frac{\sigma}{2} + \phi \left( \frac{z^3}{24\sigma^2} - \frac{z^2}{4\sigma} + \frac{z}{4} \right) \quad (4)$$

2. If  $\sigma < z \leq 2\sigma$ ,

$$\pi_1^G(z) = 1 - \phi \frac{\sigma}{2} + \phi \left( -\frac{z^3}{24\sigma^2} + \frac{z^2}{4\sigma} - \frac{z}{2} + \frac{\sigma}{3} \right) \quad (5)$$

3. If  $z > 2\sigma$ ,

$$\pi_1^G(z) = \pi_0. \quad (6)$$

Figure 3 depicts the relationship between segment distance and average division profits and illustrates the natural ambiguity in this relationship: If distance is too low there is little room to arbitrage realizations of low project-segment fit because segments have similar distances to any project; if distance is too high each division is always guaranteed to be the closest to the projects it generates, so the potential for efficient transfers also vanishes. Proposition 2 shows that optimal (static) segment distance is a simple proportion of execution-fit uncertainty  $\sigma$ , which is intuitive.

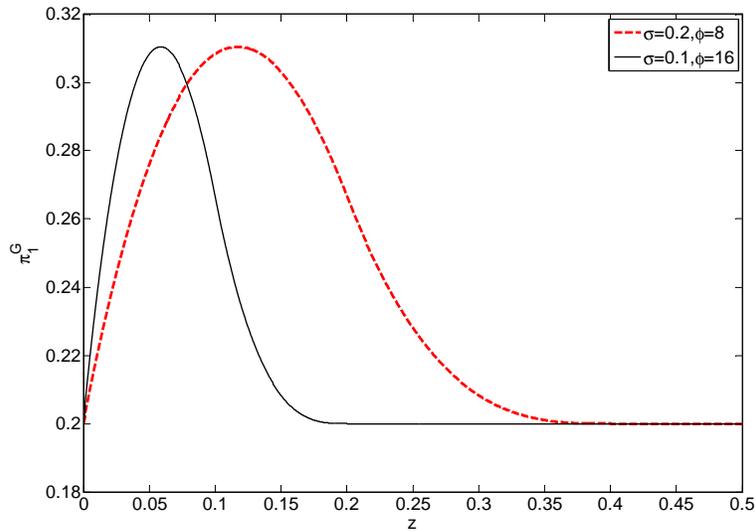
**Proposition 2** *The optimal distance between segments,  $z^*$ , is given by*

$$z^* = \sigma \left( 2 - \sqrt{2} \right), \quad (7)$$

*with associated expected BU profit of*

$$\pi_1^G(z^*) = 1 + \phi\sigma \frac{(\sqrt{2} - 4)}{6}. \quad (8)$$

If  $\sigma$  is interpreted as a measure of the inverse of specialization, then an increase in specialization (lower  $\sigma$ ) would imply that diversified firms should become more specialized too, that is, one should observe most conglomerates with segments that are closer or less diverse.



**Figure 3: Segment Distance and (Static) Profits.** The figure plots profits  $\pi_1^G$  as a function of segment distance  $z$ .

It is not clear which empirical relationship between segment distance and profits is implied by this simple static model. Inspecting figure 3, one can see that the association should be positive if most firms cluster around low segment distances; if, on the other extreme, firms are evenly distributed from 0 to 1/2, then actually the average relationship between

segment distance and value would be negative. The fact that our static model predicts this ambiguous relationship between segment distance and profits is one of the motivations for the dynamic setup we subsequently present. This ambiguity may also explain the apparent contradiction between the finance literature on corporate diversification, where relatedness is usually understood to be desirable; and the management and economic-networks literatures, who claim that economic agents spanning distant environments—“brokers”—actually draw significant rents therefrom (see Burt (2005) or Jackson (2008) for a review of these topics).

Comparing the two plots in figure 3 one observes that the relationship between segment distance and profits is *proportional* to  $\sigma$ . As long as the ratio  $\phi/\sigma$  is constant, the maximal value of synergies is the same (see proposition 2). Therefore, holding the ratio  $\phi/\sigma$  constant, it would not be possible to distinguish between an economy where  $\sigma$  is high—as in the dashed curve of figure 3—and the distribution of firms has support  $[0, 1/2]$  from an economy with low  $\sigma$ —as in the solid curve of figure 3—but the distribution of firms has support  $[0, 1/4]$ .<sup>8</sup> This point is important for our calibration, where given the argument just outlined we set  $\sigma$  at an arbitrary level.

## 3.2 Dynamics

### 3.2.1 Matching technology

We now complete our setup, by considering a dynamic continuous-time economy comprising a continuum of infinitely-lived business units (BUs) uniformly located on the circle of skill types, with a gross profit rate given by the static model developed in the previous section. For simplicity we assume that all BUs have one unit of overall resources/capacity (one project at a time in the model), and so profits and value can be understood as normalized by size. Comparing our approach to the standard neoclassical model of production, we make the assumptions that all firms have the same scale and that adjustment costs are infinite. While

---

<sup>8</sup>The caveat to this statement is that if  $\sigma$  has a clear direct empirical counterpart, then this no longer applies. This is not the case in our model.

these assumptions are extreme and unrealistic, we note however that the segment-distance effect seems robust, or at least partly invariant, to including firm size and other characteristics in the multivariate regression approach in section 2. This gives us some justification for the omission of these firm characteristics in our dynamic model.

There is an exogenous continuously-compounded discount rate denoted by  $r$  and all agents are risk-neutral. The key aspect of how we consider changes in firms' boundaries is that these happen only via merger and spin-off activity. In particular, a multi-segment firm is the product of two specialized firms that at some point in the past found it optimal to merge. For simplicity, we assume the firm splits and refocuses according to some exogenous event. We believe that modeling diversification as driven entirely by merger and spin-off activity—admittedly a stark simplification—is intuitive and not entirely unrealistic. For example, in Graham, Lemmon, and Wolf (2002) almost two thirds of the firms that increase the number of segments implement this strategy via acquisition. Also, many diversifying mergers are later divested (Ravenscraft and Scherer, 1987; Kaplan and Weisbach, 1992; Campa and Kedia, 2002; Basu, 2010).

We model mergers according to the search-and-matching models pioneered in labor economics (Diamond, 1993; Mortensen and Pissarides, 1994), an approach taken in other finance papers as well (Rhodes-Kropf and Robinson, 2008). Anjos (2010) develops a dynamic model of diversification and refocus in this spirit, which our approach also partially builds on.

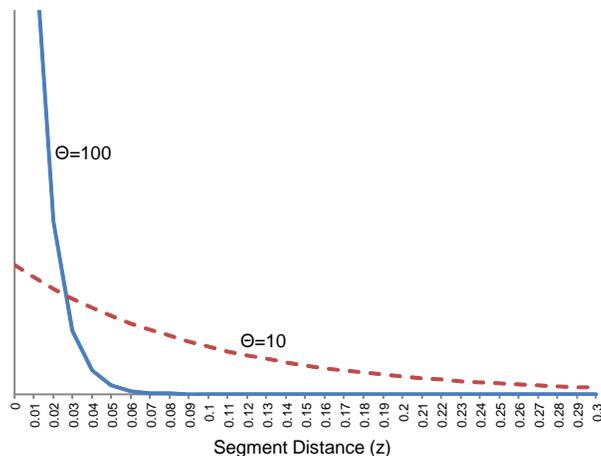
Each pair of extant specialized firms may be presented with a potential merger opportunity according to a Poisson process with intensity  $\lambda_0$ . If a meeting between two specialized firms occurs, a merger happens as long as it creates value, i.e., there are no frictions associated with the merger decision process (e.g., bargaining); surplus is shared equally across merging partners. Mergers are reversed according to an exogenous Poisson process with intensity  $\lambda_1$ .

An important ingredient of the model is how to specify the segment distance at which

matches occur. With the caveat that equilibrium has not yet been defined, if one focuses on symmetric equilibria then it makes sense that whatever technology determines the distribution of segment distance, this technology should be independent of specific locations in the circle; since all locations in the circle end up with a similar mass of business units, and moreover a similar mass of specialized and diversified firms. Based on this rationale, we specify that, conditional on a merger opportunity arising, the distance between the two segments be drawn from a truncated negative exponential with support  $[0, 1/2]$ , represented by the following density function  $m(z)$

$$m(z) = \frac{\theta e^{-\theta z}}{1 - e^{-\theta/2}}, \quad (9)$$

with  $\theta > 0$ . This density is represented in figure 4. The intuition for this distribution is that more (viable) matches tend to happen at low distances, in concert with firms being more informed about and having more connections with other firms operating in a neighborhood of their business environment. This density is also in line with the empirical distribution of firms on segment distance (qualitatively).



**Figure 4: Matching Technology – Segment Distance Distribution  $m(z)$ .** The figure plots the two density functions, with  $\theta = 10$  (dashed curve) and  $\theta = 100$  (solid curve).

The final component of the matching technology is a feedback effect. We assume that

the more conglomerates operate with distant segments, the more likely it is for far-away specialized firms to be matched. In particular we define

$$\theta \equiv \theta(\bar{z}) = \theta_0 e^{-\theta_1 \bar{z}}, \quad (10)$$

with  $\theta_0, \theta_1 > 0$ , and where  $\bar{z}$  is the average segment distance of diversified firms. Parameter  $\theta_1$  gages the extent to which  $\bar{z}$  influences the matching process. The intuition for the feedback effect is that the more conglomerates operate in somewhat directly disconnected businesses, the more the business environment of the average specialized firm also expands. It would also be natural to include the proportion of diversified firms as influencing this mechanism. However, it turns out that this would not change our main results and thus we omit this additional source of externalities (see section 4.2 below for a discussion of this issue).

### 3.2.2 The cost of headquarters

If a merger occurs at time  $t = \tau$  and is reversed at time  $t = \tau^+$ , then the profit rate of the diversified firm varies over time and is given by

$$\pi_1^G(z_\tau) - \beta_0 e^{\beta_1(t-\tau)}, \quad \forall t \in [\tau, \tau^+], \quad (11)$$

where  $\pi_1^G$  is the equilibrium profit function from the static setup, i.e. equations (4)-(5), and the second term corresponds to HQ costs. According to the expression above, HQ starts off with some cost rate  $\beta_0$ , but this cost rate increases over time, at rate  $\beta_1$ . This aims to capture that HQ becomes entrenched and more costly over time, and this is the way in which our model is able to accommodate the presence of an apparent diversification discount, in the spirit of Anjos (2010). The intuition is that most of the gains of diversification are realized when the conglomerate is young, in the form of high cash flows paid to claim holders. For this reason, the valuation of old conglomerates does not capture the value these firms created

at earlier points in time, and so an apparent discount obtains.

For simplicity we assume there are no HQ costs for specialized firms, so conglomerate HQ costs should be interpreted as the additional costs a complex diversified firm—where internal reallocation of resources is presumably taking place—entails.

### 3.3 Solving the model

First let us state the individual optimization problem. Since firms decide to merge without bilateral frictions and share merger surplus equally, the optimization problem from the perspective of business unit  $i$  is as follows:

$$J_t = \sup_{\{\tau\}} \left\{ \mathbb{E}_t \left[ \int_{u \in [t, +\infty] \cap \{\tau, \tau^+\}} e^{-r(u-t)} \left[ \pi_1^G(z_{\sup\{\tau < u\}}) - \beta_0 e^{\beta_1(u - \sup\{\tau < u\})} \right] du + \int_{u \in [t, +\infty] \setminus \{\tau, \tau^+\}} e^{-r(u-t)} \pi_0 du \right] \right\}, \quad (12)$$

where  $J_t$  is the value function of the business unit,  $\{\tau\}$  is the set of random stopping times at which the BU experiences a merger,  $\tau^+$  returns the first time after  $\tau$  at which a split takes place, and  $z_{\sup\{\tau < t\}}$  is the distance of the two divisions inside the diversified firm.

Next we define an equilibrium in this economy that is useful for our problem.

**Definition 1** (*Equilibrium*) *A Markov Perfect Equilibrium of this economy is characterized by an unchanging proportion of specialized firms  $p \in [0, 1]$ , a time-invariant merger acceptance policy  $a^*(z)$  with  $a^*(z) = 1$  if a meeting between two firms occurring at segment distance  $z$  leads to merger acceptance and  $a^*(z) = 0$  otherwise, and it is the case that the merger acceptance policy solves optimization problem (12).*

The next proposition characterizes the equilibrium value functions for specialized and diversified BUs.

**Proposition 3** *In equilibrium, the optimal policy of specialized firms is characterized by one of two mutually exclusive alternatives: (i) accepting merger matches with segment distance in an interval  $[z_L, z_H]$ , with  $0 \leq z_L \leq z_H \leq 1/2$ ; or (ii) rejecting all merger matches. In an equilibrium where some merger opportunities are desirable, the value function for a business unit inside a diversified firm,  $J_1$ , at time  $t$ , is a function of the segment distance at which the merger took place ( $z$ ) and the current duration of the merger ( $d := t - \tau$ ):*

$$J_1(z, d) = \frac{\pi_1^G(z) + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0}{r + \lambda_1 - \beta_1} e^{\beta_1 d}, \quad (13)$$

where  $J_0$  is the value of specialized firms. The value of specialized firms  $J_0$  is a constant, equal to  $\pi_0/r$  if no mergers are desirable in equilibrium. If some mergers are desirable,  $J_0$  is characterized as

$$J_0 = \frac{1}{r} \left[ \frac{r + \lambda_1}{r + \lambda_1 + \lambda_0 q} \pi_0 + \frac{\lambda_0 q}{r + \lambda_1 + \lambda_0 q} \left( \bar{\pi}_1^G - \beta_0 \frac{1}{1 - \frac{\beta_1}{r + \lambda_1}} \right) \right], \quad (14)$$

where  $q$  is the probability of merger acceptance and  $\bar{\pi}_1^G$  the average BU gross profit rate (conditional on the merger acceptance region):

$$q := \int_{z_L}^{z_H} m(z) dz \quad (15)$$

$$\bar{\pi}_1^G := \int_{z_L}^{z_H} \frac{m(z)}{q} \pi_1^G(z) dz \quad (16)$$

Equation (14) describes the equilibrium value of specialized firms, and can be interpreted as the present value of a lifetime average cash-flow rate. This average cash-flow rate is a function of the specialized cash-flow rate  $\pi_0$  and an average “time-adjusted” cash-flow rate of diversified BU’s, given by the term

$$\bar{\pi}_1^G - \beta_0 \frac{1}{1 - \frac{\beta_1}{r + \lambda_1}}.$$

The above term increases in the average gross profit of diversified BUs, and decreases in starting HQ costs  $\beta_0$  and the rate at which these costs grow  $\beta_1$ . In equation (14), what determines the weight of the specialized cash-flow rate, relative to the diversified cash-flow rate, is how frequent mergers and break-ups are, which is influenced by  $\lambda_0$ ,  $\lambda_1$ , and  $q$ . The discount rate also matters for the weighting, since as  $r$  grows the state that matters the most for value is the current one, where the firm is specialized.

It is also important to note that the value of specialized firms embeds the option value of becoming a *new* diversified firm. This point is quite relevant for our analysis, since the excess value measure does not correspond to this timing and this is why an apparent discount appears. We also note how the dynamic approach shows that using the value of specialized firms as benchmarks for determining the payoff to diversification may unjustly understate this payoff; since the value of specialized firms already reflects the option value associated with diversification gains.

Proposition 4 establishes the threshold for the efficiency gains of diversification such that mergers take place in this economy.

**Proposition 4** *There exists a threshold  $W$ , defined as*

$$W := \frac{6\beta_0}{(\sqrt{2} - 1) \left(1 - \frac{\beta_1}{r + \lambda_1}\right)}, \quad (17)$$

*such that in equilibrium  $q > 0$  if and only if  $\phi\sigma > W$ .*

The result in the proposition above shows that mergers only take place if either the location of projects is highly uncertain (high  $\sigma$ ) or the cost of BU-project misfit is high ( $\phi$ ). As derived in the static-setup section, the advantage of a conglomerate is to be able to optimize ex-post the BU-project assignment (representing resource reallocation), an option assumed to be unavailable to specialized firms. These benefits of diversification are compared to its costs, gaged by initial HQ cost  $\beta_0$ , and the rate at which this cost grows  $\beta_1$ . Interestingly,

this rate of growth is less relevant for the diversification trade-off whenever interest rates  $r$  are high, or if conglomerates break up often for exogenous reasons (high  $\lambda_1$ ). So in this model, knowing ex ante that the conglomerate will last little actually fosters diversification, which is perhaps counter-intuitive.

Propositions 5 and 6 characterize two magnitudes that are important for the calibration exercise. Proposition 3 characterizes the equilibrium proportion of diversified firms in the economy, which intuitively depends on match and break-up rates, as well as the merger acceptance probability.

**Proposition 5** *In equilibrium, the proportion of diversified firms in the economy is given by*

$$p = \frac{\lambda_1}{\lambda_1 + \lambda_0 q}, \quad (18)$$

*with  $q$  being the merger acceptance probability defined in equation (15).*

Proposition 4 further illustrates why and how a diversification discount obtains in this economy, where the average value of diversified firms may become arbitrarily negative (even though diversifying mergers are ex ante rational). The crucial determinant of the discount is the relationship between the break-up rate  $\lambda_1$  and the rate at which HQ costs grow  $\beta_1$ . For average firm values not to become too low, it needs to be the case that break-ups happen often enough early enough, such that HQ costs remain low.

**Proposition 6** *If  $\lambda_1 > \beta_1$ , the average value of diversified firms in the economy is given by*

$$E[J_1] = \frac{\bar{\pi}_1^G + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0 \lambda_1}{(\lambda_1 - \beta_1)(r + \lambda_1 - \beta_1)}, \quad (19)$$

*otherwise  $E[J_1] = -\infty$ .*

The model is solved numerically (details available from the authors), but it can be established that the equilibrium is unique.

**Proposition 7** *The equilibrium specified in definition 1 always exists and is unique.*

## 4 Results

This section addresses our main research question: Are conglomerates under-diversified in equilibrium? Our approach has two main steps. First, section 4.1 performs a baseline calibration of the dynamic model, assuming the externality channel is shut down, i.e.  $\theta_1 = 0$ . Second, and given the observed equilibrium inferred from the baseline calibration, we assess for different levels of  $\theta_1$  what the optimal outcome would be and how far the equilibrium is therefrom. Our main findings are that segment distance is on average too low compared to the optimum, although the model also implies that the efficiency loss should be low. Part of what makes these results interesting is that the theoretical model does allow for conglomerates to be over-diversified and for the externality channel to have a big impact in the differential between firm value in equilibrium and firm value at the optimum; this is shown in section 4.4. However, these alternative predictions are not borne out by data, at least in light of our calibration approach.

### 4.1 Baseline calibration

The data we use for the baseline calibration is the same used in the regressions in section 2. Summary statistics are presented in the appendix. There are a total of eight parameters to calibrate:  $r$ ,  $\theta_0$ ,  $\lambda_0$ ,  $\lambda_1$ ,  $\beta_0$ ,  $\beta_1$ ,  $\phi$ , and  $\sigma$ . As explained above we set  $\theta_1 = 0$  in this baseline calibration.

A subset of the parameters are calibrated outside the model. We set the discount rate at 10%, which seems reasonable for the average firm in the economy. We set  $\lambda_1 = 0.1$ , which implies that on average diversified firms last 10 years before splitting. While we do not have a very precise measure for this variable, Basu (2010) finds that about one third of diversifying firms reverse this decision in four years, which thus serves as a lower bound.

We perform a sensitivity analysis of our results in section 4.3, which somewhat mitigates the fact that some parameters are hard to calibrate precisely, as is the case for  $\lambda_1$ ; results are robust at least to moderate changes in parameters.

As explained in section 3.1, it would be hard to separately identify  $\sigma$ , which determines the location of positive synergies in the  $z$  space, from the location of the distribution of matches  $m(z)$ ;<sup>9</sup> so we opt to do a calibration where everything is *relative* to  $\sigma$ , which we arbitrarily set at 0.2. Finally, we use the value of specialized firms to pin down  $\phi$ . In data, the value of specialized firms (per unit of capital) has an average of 2.70. We want to obtain something that is close to this but we note that there is no cash flow growth in our model, so it seems natural to target a relatively more conservative magnitude, let us conjecture close to 2. This reasoning helps us pick  $\phi$ , which we set at 8; this yields a lower bound for the value of specialized firms of

$$\frac{\pi_0}{r} = \frac{1 - 8^{\frac{0.2}{2}}}{0.10} = 2.$$

We verify later that the option value component is not too big and that the calibrated  $J_0$  is indeed close to 2. We note that if we added constant growth to our model, say at 2% per annum, then a Tobin's  $Q$  of 2 with no growth is comparable to

$$\frac{0.1 \times 2}{0.1 - 0.02} = 2.5, \tag{20}$$

which is close to 2.70, its data counterpart.

We are left with four parameters to calibrate:  $\theta_0$ ,  $\lambda_0$ ,  $\beta_0$ , and  $\beta_1$ . We use four moments from data in order to identify these parameters. The first is the proportion of diversified firms. In our dataset, the percentage of book assets belonging to specialized firms corresponds to approximately 65% of total asset value. Thus we target  $p = 0.65$ . The other three moments we choose are the level, slope, and curvature of the relationship between excess value and

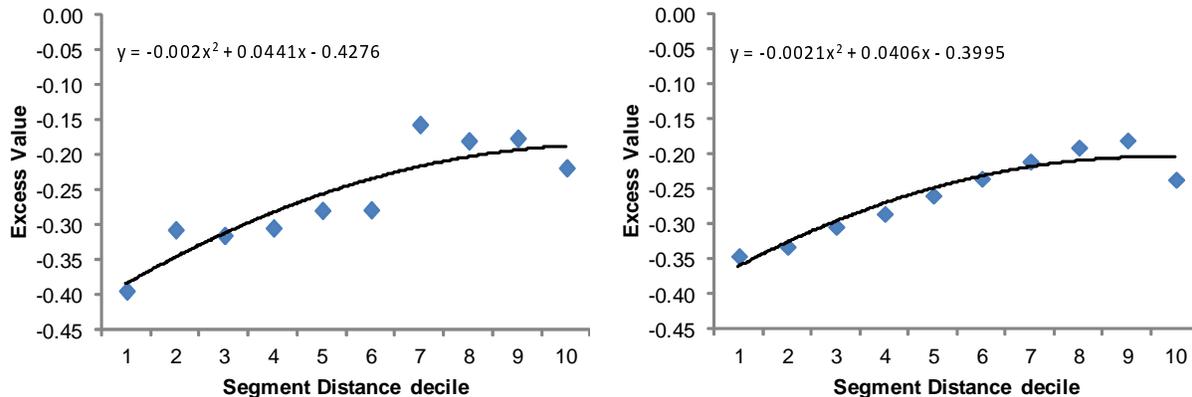
---

<sup>9</sup>See figure 3 and related text.

segment distance; in order to make data and model comparable, we define this relationship in terms of segment-distance deciles. Table 2 summarizes the choice of parameters. Our calibration is able to match data well, in the sense that we obtain  $p = 0.66$ , and figure 5 shows that we match the quadratic association between segment distance and excess value with a good fit. Just for reference, the average excess value in our data is  $-0.25$  (the apparent diversification discount), and the model generates a counterpart of  $-0.26$ .

**Table 2: Calibrated parameters.** The table shows the magnitude of each model parameter used in the baseline calibration.

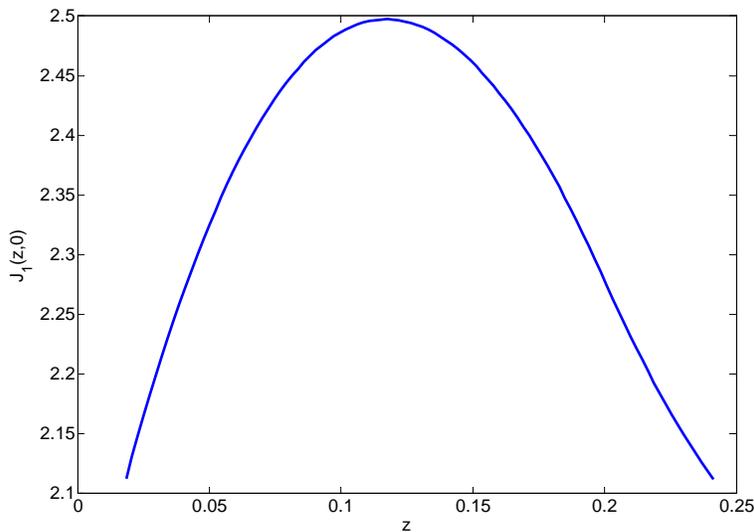
Parameter	Value
$r$	0.1
$\theta_0$	18.2
$\theta_1$	0
$\lambda_0$	0.073
$\lambda_1$	0.1
$\beta_0$	0.0125
$\beta_1$	0.0873
$\phi$	8
$\sigma$	0.2



**Figure 5: Segment Distance and Excess Value – Data (Left) and Model (Right); Baseline Calibration.** The left panel shows the empirical relationship between segment-distance deciles and excess value; excess value is defined as the log-difference between the Tobin’s  $Q$  of a conglomerate and the Tobin’s  $Q$  of a similar portfolio of specialized firms, following Berger and Ofek (1995); segment Distance is the average input-output-based distance across conglomerate segments, following Anjos and Fracassi (2011). The right panel shows the equivalent relationship in the model, under the baseline calibration.

Finally the baseline calibration delivers a value for specialized firms of  $J_0 = 2.112$ , which

implies that the value of the option to diversify is about 5% ( $0.112/2.112 - 1$ ), since the value of specialized firms without this option is simply  $\pi_0/r$ , in our case 2. If we consider a diversified firm that just obtained a match at the ideal segment distance, then corporate diversification accounts for a more sizable proportion of firm value, approximately 20% at the time the merger is executed (i.e., when HQ costs are at its lowest). This is illustrated in figure 6; where we plot the value of diversified firms at the time of the merger ( $J_1(z, 0)$ ).



**Figure 6: Conglomerate Value at Time of Merger; Baseline Calibration.** The figure plots the value of diversified firms in the merger acceptance region  $[z_L, z_H]$ , at the moment where the merger takes place.

Naturally older diversified firms have values that are below  $J_0$ , and even below 2, but that is not at all informative about the value of diversification; these firms simply already paid out whatever value they were adding from being diversified and given some frictions (reduced-form in the model, via the exogenous split) it is not simple/easy to break them apart into specialized firms. In our model refocusing adds value at some point, but management is assumed to be entrenched and such deals do not always happen. We believe this is very much in the spirit of earlier diversification literature, with the difference that it is still the case that corporate diversification is, frictions notwithstanding, an important and valuable economic activity.

## 4.2 Optimal segment distance

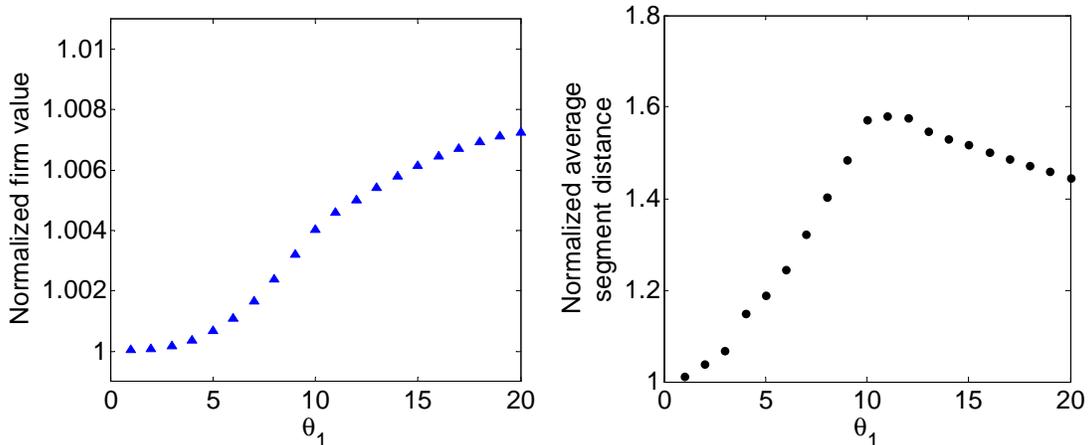
Without specific information about the extent of the externality channel, that is,  $\theta_1$  in the model, all equilibria characterized by the equation below are observationally equivalent, including the baseline calibration we just performed:

$$\theta_0 = \theta_0|_{\text{baseline}} e^{\theta_1 \bar{z}|_{\text{baseline}}}, \quad (21)$$

for all  $\theta_1 \geq 0$ . Our strategy is thus to investigate how various levels of  $\theta_1$  would affect the optimal outcome in terms of segment distance, but subject to equation (21), i.e. what is observed. For the purposes of our analysis we define optimal outcome as the *open loop* solution to the specialized firm's optimization problem (diversified firms do not make decisions in our model). This is the solution that would be implemented if all decision makers had a technology that allowed them to commit for all future periods to a particular diversification policy, here a segment-distance region where mergers are accepted. With the caveat that the solutions to our model imply a steady-state, maximizing the value of  $J_0$ , the value of specialized firms, is what best approximates the definition of maximizing ex-ante welfare; also note that increasing  $J_0$  via a different merger policy immediately implies that the value of all diversified firms is higher as well, as per equation (13). Figure 7 contains the results.

The left panel of figure 7 shows the value of specialized firms for each  $\theta_1$  scenario, normalized by the value of specialized firms in the baseline calibration. Not surprisingly, firm value increases when the externality is higher. However, the magnitude of the increase is quantitatively not very significant, always below 1%. This outcome holds in a sensitivity analysis of the calibration (see section 4.3). This result notwithstanding, we note that the theoretical model does allow for the size of the inefficiency to be relevant (see section 4.4.2); but at least in light of our data and calibration procedure, such instance does not seem to be the case.

The right panel of figure 7 shows that the optimal segment distance may be significantly



**Figure 7: Main Results.** The left panel plots the value of specialized firms,  $J_0$ , for each level of  $\theta_1$ , and normalized by the baseline-calibration magnitude of  $J_0$  (2.112). The right panel plots the average segment distance  $\bar{z}$  of the economy, for each level of  $\theta_1$ , also normalized by the baseline-calibration level of  $\bar{z}$  (0.0694).

higher than the equilibrium outcome, at least for high enough  $\theta_1$ . The upper bound for the differential occurs at an intermediate level of  $\theta_1$ , where the optimal  $\bar{z}$  is approximately 60% higher than the baseline calibration, where  $\bar{z}$  is in the order of 0.07. The non-monotonic behavior of the optimal  $\bar{z}$  occurs because when  $\theta_1$  is significantly high, increases in average segment distance may actually overshoot the ideal  $m(z)$ , in the sense that too many matches at the right tail occur; and in this region synergies no longer exist.

A final note on the results in figure 7 is that although we only show the analysis for  $\theta_1 \leq 20$ , this already implies an extremely disperse  $m(z)$ —in fact, close to uniform—for large enough  $\bar{z}$ . In that sense  $\theta_1 = 20$  is almost surely too high from an empirical perspective, since even if all conglomerates operated at maximum distance ( $z = 1/2$ ), it does not seem reasonable to assume that this would lead to uniformly distributed matches in the segment-distance space.

In the analysis above we did not consider the role that the proportion of diversified firms itself ( $1 - p$ ) could play in the externality mechanism. Indeed, it is reasonable to assume that this magnitude would also influence the matching process of specialized firms. However,  $p$  does not seem to change much (at least compared to  $\bar{z}$ ) and thus to make exposition easier we

omitted this channel. For reference,  $p$  stays within the interval  $[0.64, 0.68]$  for the scenarios of  $\theta_1$  analyzed above.

### 4.3 Sensitivity analysis of main results

This section performs a sensitivity analysis of the main result. The approach is to vary each by parameter by 20% above and below the baseline calibration.<sup>10</sup> The parameters are changed one at a time; so this exercise also works as a comparative statics analysis. Results are shown in table 3.

**Table 3: Sensitivity analysis.** The table shows one-by-one parameter changes, with a low and high scenario for each parameter, relative to the baseline calibration. In general, the high (low) scenario corresponds to (-)20% relative to the baseline, with the exception of the low scenario for  $\lambda_1$  and the high scenario for  $\beta_1$ , due to a technical restriction. The table shows the equilibrium magnitudes that obtain under each case: value of specialized firms  $J_0$ , proportion of specialized firms  $p$ , coefficients of regression of excess value on segment-distance deciles ( $C_1$  is intercept,  $C_2$  is linear coefficient,  $C_3$  is quadratic coefficient), and average segment distance  $\bar{z}$ . The table also shows the optimum  $J_0$  and  $p$  with  $\theta_1 = 10$ .

Parameter	Equilibrium						Optimum	
	$J_0$	$p$	$C_1$	$C_2$	$C_3$	$\bar{z}$	$J_0$	$\bar{z}$
$r = 0.08$	2.65	0.67	-0.38	0.04	-0.0019	0.071	2.66	0.110
$r = 0.12$	1.76	0.67	-0.41	0.04	-0.0023	0.068	1.76	0.118
$\theta_0 = 14.6$	2.12	0.65	-0.41	0.05	-0.0036	0.078	2.13	0.096
$\theta_0 = 21.8$	2.10	0.67	-0.39	0.03	-0.0009	0.062	2.11	0.118
$\lambda_0 = 0.0584$	2.09	0.70	-0.40	0.04	-0.0021	0.069	2.10	0.111
$\lambda_0 = 0.0876$	2.13	0.62	-0.40	0.04	-0.0021	0.070	2.14	0.107
$\lambda_1 = 0.0900$	2.11	0.64	-1.89	0.04	-0.0019	0.070	2.12	0.109
$\lambda_1 = 0.1200$	2.11	0.70	-0.16	0.04	-0.0020	0.068	2.11	0.111
$\beta_0 = 0.0100$	2.12	0.65	-0.33	0.04	-0.0021	0.068	2.13	0.111
$\beta_0 = 0.0150$	2.10	0.67	-0.47	0.04	-0.0021	0.071	2.11	0.112
$\beta_1 = 0.0698$	2.12	0.65	-0.15	0.04	-0.0022	0.068	2.13	0.109
$\beta_1 = 0.0950$	2.11	0.66	-1.11	0.04	-0.0020	0.070	2.12	0.109
$\phi = 6.4$	3.68	0.67	-0.22	0.02	-0.0010	0.072	3.69	0.098
$\phi = 9.6$	0.54	0.65	-1.59	0.19	-0.0099	0.068	0.55	0.113

Table 3 shows, for each scenario, what magnitudes obtain in equilibrium: the value of

<sup>10</sup>Two exceptions are the low scenario for  $\lambda_1$ , where this parameter is set at 0.09 and the high scenario for  $\beta_1$ , where this parameter is set at 0.095. The reason for these exceptions is that there is a technical restriction  $\beta_1 < \lambda_1$  (see appendix for details).

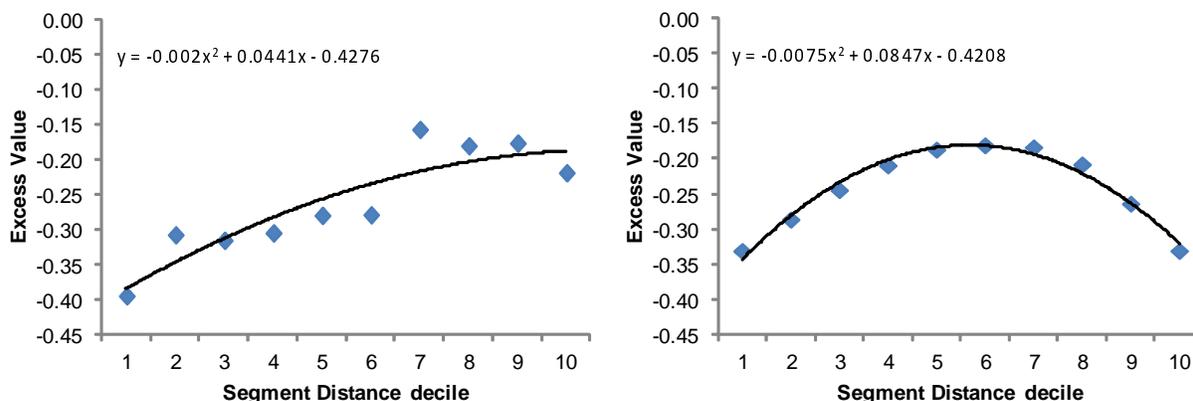
specialized firms  $J_0$ , the proportion of specialized firms  $p$ , the coefficients of the regression of excess value on segment-distance deciles shown in figure 5 ( $C_1$  is the intercept,  $C_2$  is the linear coefficient, and  $C_3$  is the quadratic coefficient), and average segment distance  $\bar{z}$ . The table also shows segment distance and  $J_0$  at the optimum, for the case where  $\theta_1 = 10$ . The main takeaway of the table in terms of our key results is that although some magnitudes do change considerably when parameters are changed, it is still the case that (i) segment distance is too low in equilibrium; and (ii) efficiency losses are small.

The table also helps understand the role each parameter plays in the model and how it is pinned down by moments in data. The risk-free rate  $r$  mainly influences equilibrium firm value, with little impact in other observable dimensions of the equilibrium. The baseline level of dispersion in firm matches, gaged by parameter  $\theta_0$ , mainly influences—at least locally—the curvature of the relationship between segment distance and excess value.  $\lambda_0$  and  $\lambda_1$ , the parameters governing the rate of matches of specialized firms and the rate of break-up of conglomerates, importantly determine the proportion of specialized firms in the economy, which is intuitive; it is also the case that  $\lambda_1$  affects the intercept of the relationship between segment distance and excess value, that is, the level of the “discount”. This is intuitive, since the growth in HQ costs over time, the mechanism behind the apparent discount, can only operate if conglomerates have a big enough lifespan.  $\beta_0$  and  $\beta_1$ , which define the initial HQ costs and the rate at which these costs grow, also influence the level of the discount as  $\lambda_0$  and  $\lambda_1$ , but have a much lower impact in terms of  $p$ , the proportion of specialized firms. Finally we can see that  $\phi$  appears as a key variable, in that it importantly affects all equilibrium magnitudes with the exception of  $p$  and  $\bar{z}$ . Nevertheless, the main results remain unchanged.

## 4.4 Alternative calibrations

### 4.4.1 Too-high segment distance

With our baseline calibration and the analysis of the role matching externalities play in the inefficiency of the equilibrium, we found that the observed average segment distance was too low. A natural question is whether the model ever predicts distance to be too high, i.e., that conglomerates are over-diversified. This can indeed be the case when the natural segment-distance matching function is already too disperse. We took the baseline calibration and changed only two parameters:  $\sigma$  is now reduced to 0.05 (so positive synergies are now located closer to  $z = 0$ ), and  $\phi$  is increased to 32 (to keep the value of synergies similar). We denote this calibration as “alternative calibration 1”.



**Figure 8: Segment Distance and Excess Value – Data (Left) and Model (Right); Alternative Calibration 1.** The left panel shows the empirical relationship between segment-distance deciles and excess value; excess value is defined as the log-difference between the Tobin’s  $Q$  of a conglomerate and the Tobin’s  $Q$  of a similar portfolio of specialized firms, following Berger and Ofek (1995); segment Distance is the average input-output-based distance across conglomerate segments, following Anjos and Fracassi (2011). The right panel shows the equivalent relationship in the model, under alternative calibration 1. All parameters are set at the level of the baseline calibration, except  $\sigma = 0.05$  and  $\phi = 32$ .

Whereas the value of specialized firms is kept at a level close to 2 as in the baseline calibration (more precisely, at approx. 2.11), and the proportion of specialized firms is close to 0.65 (more precisely, at approx. 0.70), figure 8 shows that the model predicts a negative parabola in terms of the relationship between segment distance and excess value, unlike we observed in data. The intuition is that for this alternative calibration, matches are occurring

with high likelihood at all segment distances where synergies are positive, and thus one observes the full spectrum of the synergy-generating function (instead of mostly the left part, the case in the baseline calibration). It is important to note that many matches are actually occurring in the region where no synergies exist; the merger acceptance probability  $q$  drops to 0.59 (it is 0.70 in the baseline calibration).

For this calibration, the optimal segment distance is now lower than what obtains in equilibrium, that is, the social planner uses the externality channel to make more matches happen in the positive-synergy region. Whereas in equilibrium we have  $\bar{z} = 0.028$ , for  $\theta_1 = 20$  the optimal is  $\bar{z} = 0.024$ ; i.e. a decrease of almost 15%.

#### 4.4.2 Externalities could matter

This section shows that although the baseline calibration predicts quantitatively unimportant efficiency losses, the model does allow for the possibility of these being large. We changed the following parameters:  $\theta_0$  with  $\theta_1 = 0$  is now set at 4 times its level in the baseline calibration;  $\lambda_0$  is also set at 4 times its level of the baseline calibration. We denote this calibration as “alternative calibration 2”. In equilibrium, this leads to  $J_0 = 2.084$ , whereas the optimum with  $\theta_1 = 20$  leads to  $J_0 = 2.218$ , an increase in value of more than 6%. What rules out alternative calibration 2 is that the predicted segment-distance/excess-value relationship does not exhibit a good fit with data.

## 5 Conclusion

This paper develops a novel approach to corporate diversification, where synergies are explicitly modeled as function of how far away conglomerate segments are, in an industry-spatial sense. Corporate diversification takes place via merger activity and, also, the presence of conglomerates in the economy influences the merger-opportunity generating process; the more diversified firms there are, the more likely it is that far-away specialized firms are

matched and presented with the opportunity to merge. Combining data and our dynamic setup we find that it is likely that US conglomerates are under-diversified relative to the optimal outcome, although efficiency losses seem low. Finally, our model implies that corporate diversification is an important and valuable economic activity, and that it accounts for a sizable proportion of firm value.

# A Appendix

## A.1 Proofs

### Proof of proposition 1.

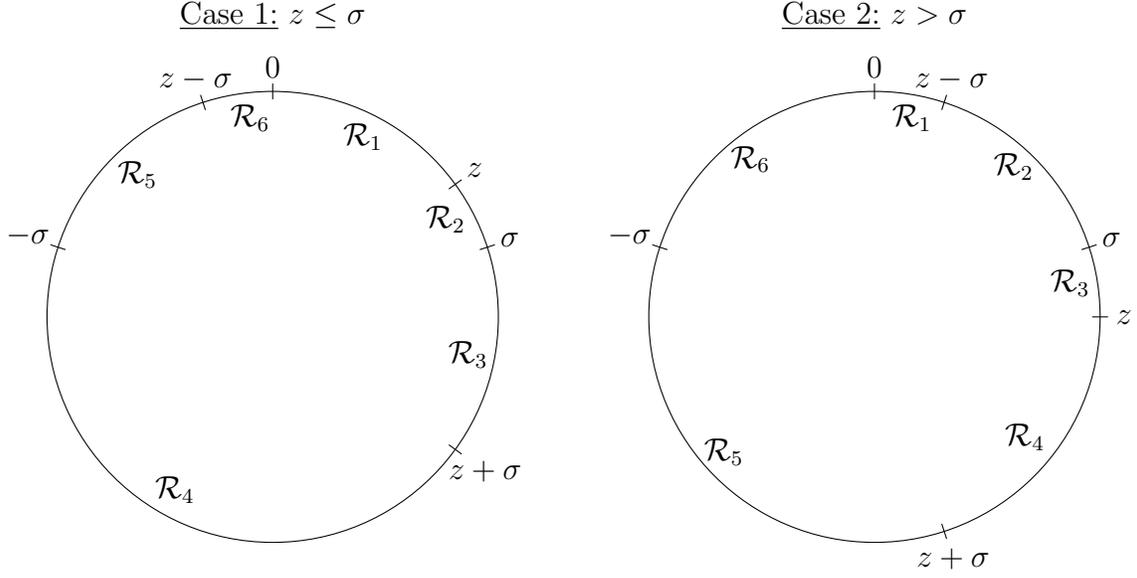
First let us set, without loss of generality,  $\alpha_i = 0$  and  $\alpha_j < 1/2$ ; also recall that we are assuming  $\sigma < 1/4$ . It may additionally be useful to clarify the convention we are employing with respect to circle location, namely that  $N_1 + x$  is equivalent to  $N_2 + x$ , for any two integers  $N_1$  and  $N_2$ , and all  $x \in [0, 1]$ .

#### Case 1: $z \leq \sigma$

Consider the left circle in figure A.1. Let us denote the six adjacent regions in the following way. Starting at 0 and going clockwise until  $z$  defines region  $\mathcal{R}_1$ ; starting at  $z$  and going clockwise until  $\sigma$  defines region  $\mathcal{R}_2$ ; and so forth. The location of the project generated by  $i$  can occur in regions 1, 2, 5, or 6; the location of the project generated by  $j$  can occur in regions 1, 2, 3, or 6. Since profits are linear in distance between BUs and projects, the optimal allocation of execution is the one that minimizes total “travel” from the (assigned) projects to each division/BU. Inspection of the different possibilities allows us to determine the optimal policy for each case, with results shown in table A.1.

Let us take the perspective of BU  $i$  and define  $E[z_{i,P_i^*}]$  as the expected distance of  $\alpha_i$  to the project optimally undertaken by  $i$ . This can be written as

$$\begin{aligned}
 E[z_{i,P_i^*}] &= \\
 &= \Pr\{\alpha_{P_i} \in \mathcal{R}_1\} \left[ \Pr\{\alpha_{P_j} \in \mathcal{R}_1\} E[\min(z_{i,P_i}, z_{i,P_j}) | \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_1] + \right. \\
 &\quad \left. + \Pr\{\alpha_{P_j} \in \mathcal{R}_6\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_6] + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}) E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_1] \right] + \\
 &\quad + \Pr\{\alpha_{P_i} \in \mathcal{R}_2\} \left[ \Pr\{\alpha_{P_j} \in \mathcal{R}_1\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_1] + \right. \\
 &\quad \left. + \Pr\{\alpha_{P_j} \in \mathcal{R}_6\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_6] + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}) E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_2] \right] +
 \end{aligned}$$



**Figure A.1:** Splitting the circle into regions. In the left example,  $\sigma = 0.2$  and  $z = 0.15$ . In the right example,  $\sigma = 0.2$  and  $z = 0.25$ .

$$+ \Pr\{\alpha_{P_i} \in \mathcal{R}_5\}E[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_5] + \Pr\{\alpha_{P_i} \in \mathcal{R}_6\}E[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_6]. \quad (\text{A.1})$$

The expression (as a function of parameters) of each of the components in equation (A.1) is presented in table A.2.

We are omitting the explicit integration procedures, since all conditional distributions are uniform (in the relevant region), so probabilities and expected distances are generally simple functions of (region) arc length; the slightly more complex case is the computation of  $E[\min(z_{i,P_i}, z_{j,P_j})|\dots]$ , where we used a standard result on order statistics for random variables drawn from independent uniform distributions.<sup>A.1</sup>

Inserting the expressions from table A.2 into equation (A.1), and after a few steps of algebra,

<sup>A.1</sup>The expected value of the  $k$ -th order statistic for a sequence of  $n$  independent uniform random variables on the unit interval is given by

$$\frac{k}{n+k}.$$

In our case,  $k = 1$  and  $n = 2$  (the two projects), and the random variables have support  $[0, z]$ , which yields  $E[\min(z_{i,P_i}, z_{j,P_j})|\dots] = z/3$ .

Location of $\alpha_{P_i}$	Location of $\alpha_{P_j}$	Optimal allocation policy
$\mathcal{R}_1$	$\mathcal{R}_1$	Swap if and only if $\alpha_{P_j} < \alpha_{P_i}$ .
$\mathcal{R}_1$	$\mathcal{R}_2$	Never swap.
$\mathcal{R}_1$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_1$	$\mathcal{R}_6$	Always swap.
$\mathcal{R}_2$	$\mathcal{R}_1$	Always swap.
$\mathcal{R}_2$	$\mathcal{R}_2$	Indifferent (no swap assumed).
$\mathcal{R}_2$	$\mathcal{R}_3$	Indifferent (no swap assumed).
$\mathcal{R}_2$	$\mathcal{R}_6$	Always swap.
$\mathcal{R}_5$	$\mathcal{R}_1$	Never swap.
$\mathcal{R}_5$	$\mathcal{R}_2$	Never swap.
$\mathcal{R}_5$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_5$	$\mathcal{R}_6$	Indifferent (no swap assumed).
$\mathcal{R}_6$	$\mathcal{R}_1$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_2$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_6$	Indifferent (no swap assumed).

**Table A.1:** Optimal allocation policy (swap/no-swap) when two projects occur, as a function of project location; with  $z \leq \sigma$ .

one obtains

$$E [z_{i,P_i^*}] = \frac{1}{24\sigma^2} (-z^3 + 6\sigma z^2 - 6\sigma^2 z + 12\sigma^3), \quad (\text{A.2})$$

which implies equation (4) in the proposition.

Case 2:  $z > \sigma$

For this case let us make the additional assumption that  $z \leq 2\sigma$ . This assumption is made without loss of generality, since for  $z > 2\sigma$  there cannot be any gains from diversification and the two-division conglomerate is simply a collection of two specialized business units, each undertaking its own projects (this corresponds to equation (6) in the proposition). Let us again partition the circle into six regions, depicted in the right of figure A.1. Similarly as in the previous case, we define region  $\mathcal{R}_1$  as the arc between 0 and  $z - \sigma$ , region  $\mathcal{R}_2$  as the arc between  $z - \sigma$  and  $\sigma$ , and so on. The location of the project generated by  $i$  can occur in regions 1, 2, or 3; the location of the project generated by  $j$  can occur in region 2, 3, or 4. Table A.3 shows the optimal execution policy for each scenario.

Item	Expression
$\Pr\{\alpha_{P_i} \in \mathcal{R}_1\}$	$\frac{z}{2\sigma}$
$\Pr\{\alpha_{P_j} \in \mathcal{R}_1\}$	$\frac{z}{2\sigma}$
$E[\min(z_{i,P_i}, z_{j,P_j})   \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_1]$	$\frac{z}{3}$
$\Pr\{\alpha_{P_j} \in \mathcal{R}_6\}$	$\frac{\sigma-z}{2\sigma}$
$E[z_{i,P_j}   \alpha_{P_j} \in \mathcal{R}_6]$	$\frac{\sigma-z}{2}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_1]$	$\frac{z}{2}$
$\Pr\{\alpha_{P_i} \in \mathcal{R}_2\}$	$\frac{\sigma-z}{2\sigma}$
$E[z_{i,P_j}   \alpha_{P_j} \in \mathcal{R}_1]$	$\frac{z}{2}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_2]$	$\frac{z+\sigma}{2}$
$\Pr\{\alpha_{P_i} \in \mathcal{R}_5\}$	$\frac{z}{2\sigma}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_5]$	$\frac{2\sigma-z}{2}$
$\Pr\{\alpha_{P_i} \in \mathcal{R}_6\}$	$\frac{\sigma-z}{2\sigma}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_6]$	$\frac{\sigma-z}{2}$

**Table A.2:** Auxiliary table for derivation of equation (A.2).

Again let us take the position of BU  $i$ ; we can then write

$$\begin{aligned}
E[z_{i,P_i^*}] &= \\
&= \Pr\{\alpha_{P_i} \in \mathcal{R}_1\}E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_1] + \Pr\{\alpha_{P_i} \in \mathcal{R}_6\}E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_6] \\
&+ \Pr\{\alpha_{P_i} \in \mathcal{R}_2\} \left[ \Pr\{\alpha_{P_j} \in \mathcal{R}_2\}E[\min(z_{i,P_i}, z_{i,P_j}) | \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_2] + \right. \\
&\left. + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_2\})E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_2] \right]. \tag{A.3}
\end{aligned}$$

The expression of each of the components in equation (A.3) is presented in table A.4.

Inserting the expressions from table A.4 into equation (A.3), and after a few steps of algebra, one obtains

$$E[z_{i,P_i^*}] = \frac{1}{24\sigma^2} (z^3 - 6\sigma z^2 + 12\sigma^2 z + 4\sigma^3), \tag{A.4}$$

Location of $\alpha_{P_i}$	Location of $\alpha_{P_j}$	Optimal allocation policy
$\mathcal{R}_1$	$\mathcal{R}_2$	Never swap.
$\mathcal{R}_1$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_1$	$\mathcal{R}_4$	Never swap.
$\mathcal{R}_2$	$\mathcal{R}_2$	Swap if and only if $\alpha_{P_j} < \alpha_{P_i}$ .
$\mathcal{R}_2$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_2$	$\mathcal{R}_4$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_2$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_3$	Never swap.
$\mathcal{R}_6$	$\mathcal{R}_4$	Never swap.

**Table A.3:** Optimal allocation policy (swap/no-swap) when two projects occur, as a function of project location; with  $z > \sigma$ .

Item	Expression
$\Pr\{\alpha_{P_i} \in \mathcal{R}_1\}$	$\frac{z-\sigma}{2\sigma}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_1]$	$\frac{z-\sigma}{2}$
$\Pr\{\alpha_{P_i} \in \mathcal{R}_6\}$	$\frac{1}{2}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_6]$	$\frac{\sigma}{2}$
$\Pr\{\alpha_{P_i} \in \mathcal{R}_2\}$	$\frac{2\sigma-z}{2\sigma}$
$\Pr\{\alpha_{P_j} \in \mathcal{R}_2\}$	$\frac{2\sigma-z}{2\sigma}$
$E[\min(z_{i,P_i}, z_{j,P_j})   \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_2]$	$\frac{2z-\sigma}{3}$
$E[z_{i,P_i}   \alpha_{P_i} \in \mathcal{R}_2]$	$\frac{z}{2}$

**Table A.4:** Auxiliary table for derivation of equation (A.4).

which implies expression (5) in the proposition. ■

### Proof of proposition 2.

Let us start by conjecturing that the optimal segment distance is smaller than  $\sigma$ . Then we need to obtain the first-order condition with respect to equation (4), which is

$$\frac{z^2}{8\sigma^2} - \frac{z}{2\sigma} + \frac{1}{4} = 0 \Leftrightarrow z^2 - 4z\sigma + 2\sigma^2 = 0.$$

The two roots of the above quadratic are given by, after a few steps of algebra,

$$z = \sigma \left( 2 \pm \sqrt{2} \right).$$

The root with the plus sign before the square root term cannot be a solution, since it would imply  $z^* \geq 2\sigma$ . Therefore we are left with the other root, i.e. equation (7) in the proposition. The next step in the proof is to verify our initial conjecture that the optimal  $z$  cannot lie in the second branch of the value function. To prove this, it is sufficient to show that equation (5) is never upward-sloping in its domain:

$$-\frac{z^2}{8\sigma^2} + \frac{z}{2\sigma} - \frac{1}{2} \leq 0 \Leftrightarrow z^2 - 4\sigma z + 4\sigma^2 \geq 0 \Leftrightarrow (z - 2\sigma)^2 \geq 0,$$

which concludes the proof. ■

### Proof of proposition 3.

[*Note: To understand the derivations below, it may be useful to recall that a random variable following a Poisson process with intensity  $x$  is realized over the next time infinitesimal  $dt$  with probability  $x dt$ .*]

We focus on the equilibrium where mergers take place in equilibrium (the other case is trivial). The solution to the firm's optimization problem (12) is a simple application in real options theory, where the exercise threshold corresponds to a minimum level for the cash-flow rate of a diversified BU. This minimum cash-flow rate maps onto a region  $[z_L, z_H]$  around the static optimum  $z^*$ ; and where in particular  $\pi_1^G(z_L) = \pi_1^G(z_H)$ .

The solution to the problem described in expression (12), given financial markets' equilibrium, needs to verify the following conditions:

$$\left\{ \begin{array}{l} rJ_1(z, t, \tau) dt = [\pi_1^G(z) - \beta_0 e^{\beta_1(t-\tau)}] dt + E_t[dJ_t], \end{array} \right. \quad (\text{A.5})$$

$$\left\{ \begin{array}{l} rJ_0 dt = \pi_0 dt + E_t[dJ_t] \end{array} \right. \quad (\text{A.6})$$

Equation (A.5) can be transformed into an ordinary differential equation (and where for notational simplicity we set  $\tau = 0$ ):

$$\begin{aligned} rJ_1(z, t) dt &= \pi_1^G(z) dt + \lambda_1 dt [J_0 - J_1(z, t)] + (1 - \lambda_1 dt) \frac{\partial J_1(z, t)}{\partial t} dt \Leftrightarrow \\ J_1(z, t)(r + \lambda_1) - \frac{\partial J_1(z, t)}{\partial t} + \beta_0 e^{\beta_1 t} - (\pi_1^G(z) + \lambda_1 J_0) &= 0 \end{aligned} \quad (\text{A.7})$$

The economically-meaningful solution for the differential equation takes the form

$$J_1(z, t) = A_1 + A_2 e^{\beta_1 t}, \quad (\text{A.8})$$

where  $A_1$  and  $A_2$  are constants. Using expression (A.8) and inserting it into the differential equation (A.7), one easily pins down  $A_1$  and  $A_2$ :

$$\begin{aligned} A_1 &= \frac{-\beta_0}{r + \lambda_1 - \beta_1} \\ A_2 &= \frac{r \left( \frac{\pi_1^G(z)}{r} \right) + \lambda_1 J_0}{r + \lambda_1} \end{aligned}$$

This completes the derivation of the expression for  $J_1$  in the proposition.

Equation A.6 can also be expressed as a (trivial) functional equation:

$$\begin{aligned} rJ_0 dt &= \pi_0 dt + \lambda_0 dt q \{E[J_1(z, t + dt)|z \in [\underline{z}, \bar{z}]] - J_0\} + (1 - \lambda_0 dt q) 0 dt \Leftrightarrow \\ J_0(r + \lambda_0 q) &= \pi_0 + \lambda_0 q \{E[J_1(z, t + dt)|z \in [\underline{z}, \bar{z}]]\} \Leftrightarrow \\ J_0 &= \frac{1}{r + \lambda_0 q} (\pi_0 + \lambda_0 q \{E[J_1(z, t + dt)|z \in [\underline{z}, \bar{z}]]\}) \end{aligned} \quad (\text{A.9})$$

where  $q$  is the probability of merger acceptance, conditional on a match taking place, defined in the proposition. Combining equation (A.9) and equation (13) yields expression (14). ■

**Proof of proposition 4.**

Let us start with the sufficiency argument. If  $q = 0$  then no specialized firm ever wants to merge, even in the best possible case, i.e. a match where  $z = z^*$ . We also know that in this economy  $J_0 = \pi_0/r$ . Combining this with the optimality of the decision not to merge in the best possible case, we have the following condition:

$$J_1(z^*, 0) \leq \frac{\pi_0}{r} \Leftrightarrow \frac{\pi_1^G(z^*) + \lambda_1 \frac{\pi_0}{r}}{r + \lambda_1} - \frac{\beta_0}{r + \lambda_1 - \beta_1} \leq \frac{\pi_0}{r},$$

where we used equation (13). Replacing  $\pi_0$  and  $\pi_1^G(z^*)$  by their expressions as a function of primitives  $\sigma$  and  $\phi$  (equations (3) and (8)); and after a few steps of algebra, yields the result  $\phi\sigma \leq W$ . For the necessity part of the proof we note that  $q = 0$  could not be an equilibrium if  $\phi\sigma > W$ , since, by the argument above, there would be some mergers worth executing (which is inconsistent with  $q = 0$ ). Since, by proposition 7 an equilibrium always exists, it must be the case that  $q > 0$  holds in equilibrium. ■

### Proof of proposition 5.

Since in equilibrium the distribution of firms is stationary, it needs to be the case that the mass of specialized firms becoming diversified over an infinitesimal  $dt$ ,  $\lambda_0 q dt$ , be the same as the mass of firms refocusing, which is  $(1 - p)\lambda_1 dt$ . Simplification of this equality yields the expression in the proposition. ■

### Proof of proposition 6.

For all diversified firms operating at segment distance  $z$ , the distribution of ages follows a negative exponential distribution with parameter  $\lambda_1$  (the break-up rate of conglomerates). It follows that

$$\begin{aligned} \mathbb{E}[J_1(z, d)|z = \tilde{z}] &= \frac{\pi_1^G(\tilde{z}) + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0}{r + \lambda_1 - \beta_1} \int_0^\infty e^{\beta_1 s} \lambda_1 e^{-\lambda_1 s} ds = \\ &= \frac{\pi_1^G(\tilde{z}) + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0 \lambda_1}{(\lambda_1 - \beta_1)(r + \lambda_1 - \beta_1)}. \end{aligned}$$

Integrating over all possible  $\tilde{z}$  yields the result in the proposition. ■

**Proof of proposition 7.**

First note that the equilibrium exists and is unique for  $\phi\sigma \leq W$ , where  $W$  is defined in proposition 4. In this simple equilibrium, irrespective of starting history with some conglomerates or not, the steady state comprises all firms being specialized (i.e.  $p = 1$ ). Next let us establish that an equilibrium always exists for  $\phi\sigma > W$ . The optimality condition that needs to be verified is given by the standard dynamic-programming principle:

$$a^*(z) = 1 \Leftrightarrow J_1(z, d = 0) \geq J_0,$$

which implies equilibrium  $z_L$  and  $z_H$  such that

$$J_1(z_L, 0) = J_1(z_H, 0) = J_0.$$

Using expressions (14) and (13), further manipulation yields (focusing on  $z_L$  without loss of generality)

$$\begin{aligned} \frac{\pi_1^G(z_L)}{r + \lambda_1} - \frac{\beta_0}{r + \lambda_1 - \beta_1} &= J_0 \left( 1 - \frac{\lambda_1}{r + \lambda_1} \right) \Leftrightarrow \pi_1^G(z_L) - \frac{\beta_0}{1 - \frac{\beta_1}{r + \lambda_1}} = J_0 r \Leftrightarrow \\ \pi_1^G(z_L) - \frac{\beta_0}{1 - \frac{\beta_1}{r + \lambda_1}} &= \frac{r + \lambda_1}{r + \lambda_1 + \lambda_0 q} \pi_0 + \frac{\lambda_0 q}{r + \lambda_1 + \lambda_0 q} \left( \pi_1^G - \beta_0 \frac{1}{1 - \frac{\beta_1}{r + \lambda_1}} \right) \end{aligned} \quad (\text{A.10})$$

One can always find a  $z_0$  such that

$$\pi_1^G(z_0) - \frac{\beta_0}{1 - \frac{\beta_1}{r + \lambda_1}} = \pi_0,$$

since we are analyzing the case  $\phi\sigma > W$ . Then we can write (A.10) as

$$\pi_1^G(z_L) - \frac{\beta_0}{1 - \frac{\beta_1}{r + \lambda_1}} = \frac{r + \lambda_1}{r + \lambda_1 + \lambda_0 q} \left( \pi_1^G(z_0) - \beta_0 \frac{1}{1 - \frac{\beta_1}{r + \lambda_1}} \right) +$$

$$\begin{aligned}
& + \frac{\lambda_0 q}{r + \lambda_1 + \lambda_0 q} \left( \bar{\pi}_1^G - \beta_0 \frac{1}{1 - \frac{\beta_1}{r + \lambda_1}} \right) \Leftrightarrow \\
\pi_1^G(z_L) & = \frac{r + \lambda_1}{r + \lambda_1 + \lambda_0 q} \pi_1^G(z_0) + \frac{\lambda_0 q}{r + \lambda_1 + \lambda_0 q} \bar{\pi}_1^G \quad (\text{A.11})
\end{aligned}$$

Noting that  $\pi_1^G(z_L) \leq \bar{\pi}_1^G \leq \pi_1^G(z^*)$ , then continuity implies the existence of  $z_L \in [z_0, z^*]$  that satisfies equation (A.11). Continuity also implies the existence of  $z_H$  such that  $\pi_1^G(z_H) = \pi_1^G(z_L)$ . Uniqueness follows from continuity and the fact that the equilibrium is unique at  $\phi\sigma \leq W$ . ■

## A.2 Definitions of variables

- *Capital Expenditures (CAPEX)*: Funds used for additions to PP&E, excluding amounts arising from acquisitions (Source: CAPEX variable in COMPUSTAT).
- *Earnings Before Interest and Taxes (EBIT)*: Net Sales, minus Cost of Goods Sold minus Selling, General & Administrative Expenses minus Depreciation and Amortization (Source: EBIT variable in COMPUSTAT).
- *Excess Value*: The log-difference between the Tobin's  $Q$  of a conglomerate and the assets-weighted Tobin's  $Q$  of a similar portfolio of specialized firms, using the 6-digit Input-Output industry classification system (Source: CRSP, COMPUSTAT, BEA, and Authors' Calculations).
- *Number of Segments*: The number of unique segments of a conglomerate using the 6-digit Input-Output industry classification system (Source: COMPUSTAT SEGMENTS and BEA).
- *Related Segments*: The number of unique segments of a conglomerate using the 6-digit Input-Output industry classification system, minus the number of unique segments of a conglomerate using the 3-digit Input-Output industry classification system, defined

as in Berger and Ofek (1995) (Source: COMPUSTAT SEGMENTS and BEA).

- *Return on Assets* (ROA): The sum of current Income Before Extraordinary Items (IB) and current Total Interest and related Expenses (XINT), divided by the value of Total Assets (AT) at the beginning of the year (Source: COMPUSTAT).
- *Sales*: Gross sales reduced by cash discounts, trade discounts, and returned sales (Source: SALE variable in COMPUSTAT).
- *Segment Distance*: the distance between any two industries the conglomerate participates in, averaged across all pairs (Source: COMPUSTAT SEGMENTS, BEA, and Authors' Calculations).
- *Tobin's Q*: The sum of total assets (AT) minus the book value of equity (BE) plus the market capitalization (Stock Price at the end of the year (PRCC\_F) times the number of shares outstanding (CSHO)), divided by the total assets (AT) (Source: COMPUSTAT).
- *Total Assets*: The total assets of a company (Source: AT variable in COMPUSTAT).
- *Vertical Relatedness*: Constructed following Fan and Lang (2000). Measures the average input-output flow intensity (in Anjos and Fracassi (2011)) between each of the conglomerate's non-primary segments and the conglomerate's primary segment; averaged over all non-primary segments. (Source: COMPUSTAT SEGMENTS, BEA, and Authors' Calculations).

### A.3 Summary statistics of dataset

**Table A.5: Summary Statistics.** The table presents means, standard deviations, percentiles, and the number of observations for each variable. Excess Value is defined as the log-difference between the Tobin's  $Q$  of a conglomerate and the Tobin's  $Q$  of a similar portfolio of specialized firms, following Berger and Ofek (1995). Segment Distance is defined as the average distance for all possible pairs of segments participated by the conglomerate, using the industry distance definition from Anjos and Fracassi (2011). N. Segments is the number of segments in the conglomerate, using the 6-digit Input-Output industry classification system. Related Segments is measured as the difference between the number of segments of a conglomerate using the 6-digit Input-Output industry classification system and the unique number of segments using the 3-digit Input-Output industry classification system. All variables are defined in detail in the appendix.

**Panel A: Conglomerates' Data**

	Mean	Median	St.Dev.	5-th Perc.	95-th Perc.	Obs.
Tobin's $Q$	1.64	1.24	1.63	0.78	3.49	7,947
Excess Value	-0.25	-0.26	0.65	-1.25	0.82	7,892
Segment Distance	9.78	7.34	9.82	0.27	27.00	7,926
N. Segments	2.53	2.00	0.85	2.00	4.00	7,926
Related Segments	0.26	0.00	0.53	0.00	1.00	7,926
EBIT/Sales	-0.14	0.06	5.95	-0.37	0.23	7,734
CAPEX/Sales	0.13	0.05	0.99	0.01	0.39	7,854
Total Assets (\$MM)	3,484	400	11,683	8	16,462	7,947

**Panel B: Single-Segment Firms' Data**

	Mean	Median	St.Dev.	5-th Perc.	95-th Perc.	Obs.
Tobin's $Q$	2.70	1.64	3.35	0.77	8.03	69,729
EBIT/Sales	-6.29	0.03	170.80	-5.42	0.27	64,393
CAPEX/Sales	1.09	0.04	45.57	0.00	0.81	68,689
Total Assets (\$MM)	742	67	3,585	2	2,682	69,729

## References

- Ahern, Kenneth R., and Jarrad Harford, 2012, The importance of industry links in merger waves, Forthcoming, *Journal of Finance*.
- Anjos, Fernando, 2010, Costly refocusing, the diversification discount, and the pervasiveness of diversified firms, *Journal of Corporate Finance* 16, 276–287.
- Anjos, Fernando, and Cesare Fracassi, 2011, Shopping for information? Diversification and the network of industries, Working paper, available at SSRN.
- Basu, Nilanjan, 2010, Trends in corporate diversification, *Financial Markets and Portfolio Management* 24, 87–102.
- Berger, Philip G., and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–65.
- Burt, Ronald S., 2005, *An Introduction to Social Capital* (Oxford University Press).
- Campa, José Manuel, and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 1731–1762.
- Chou, Eric S., 2007, The boundaries of firms as information barriers, *RAND Journal of Economics* 38, 733–746.
- Custódio, Cláudia, 2009, Mergers & acquisitions accounting can explain the diversification discount, working paper, London School of Economics.
- Diamond, Peter A., 1993, Search, sticky prices, and inflation, *Review of Economic Studies* 60, 53–68.
- Fan, Joseph, and Larry Lang, 2000, The measurement of relatedness: An application to corporate diversification, *Journal of Business* 73, 629–60.

- Gomes, Joao, and Dmitry Livdan, 2004, Optimal diversification: reconciling theory and evidence, *Journal of Finance* 59, 505–535.
- Graham, John. R., Michael L. Lemmon, and Jack G. Wolf, 2002, Does corporate diversification destroy value?, *Journal of Finance* 57, 695–719.
- Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.
- Jackson, Matthew O., 2008, *Social and Economic Networks* (Princeton University Press).
- Kaplan, S., and M. S. Weisbach, 1992, The success of acquisitions: evidence from divestitures, *Journal of Finance* 47, 107–138.
- Lang, Larry H. P., and René M. Stulz, 1994, Tobin’s  $q$ , corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248–1280.
- Mortensen, Dale T., and Christopher A. Pissarides, 1994, Job creation and job destruction in the theory of unemployment, *Review of Economic Studies* 61, 397–415.
- Ravenscraft, D. J., and F. M. Scherer, 1987, *Mergers, sell-offs, and economic efficiency* (Brookings Institution).
- Rhodes-Kropf, Matthew, and David T. Robinson, 2008, The market for mergers and the boundaries of the firm, *Journal of Finance* 63, 1169–1211.
- Scharfstein, David S., and Jeremy C. Stein, 2000, The dark side of internal capital markets: divisional rent-seeking and inefficient investment, *Journal of Finance* 55, 2537–2564.
- Stein, Jeremy C., 1997, Internal capital markets and the competition for corporate resources, *Journal of Finance* 52, 111–133.

Tate, Geoffrey, and Liu Yang, 2011, The bright side of corporate diversification: Evidence from internal labor markets, Working paper, available at SSRN.

Villalonga, Belén, 2004, Diversification discount or premium? New evidence from the business information tracking series, *Journal of Finance* 59, 479–506.