Technological Specialization and Corporate Diversification*

Fernando Anjos[†]　　Cesare Fracassi[‡]

Abstract

We develop a dynamic model of diversifying mergers, with three key features: (i) synergies from combining complementary technologies are affected by an exogenous level of technological specialization; (ii) conglomerates incur a marginal cost associated with organizational complexity; (iii) mergers are modeled as a search-and-matching process. A calibrated version of the model simultaneously matches three corporate-diversification magnitudes: the proportion of assets allocated to conglomerates; the diversification discount; and a positive association between conglomerate value and segment distance, an industry-network measure of technological complementarity. The calibrated model also shows how an increase in technological specialization can account for empirical corporate-diversification trends.

December 21, 2013

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

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

*The authors thank comments from and discussions with Andres Almazan, Aydoğan Alti, 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 and 2013 North American Summer meetings of the Econometric Society.

[†]University of Texas at Austin, McCombs School of Business, 2110 Speedway, Stop B6600, Austin TX 78712. Telephone: (512) 232-6825. E-mail: fernando.anjos@mccombs.utexas.edu

[‡]University of Texas at Austin, McCombs School of Business, 2110 Speedway, Stop B6600, Austin TX 78712. Telephone: (512) 232-6843. E-mail: cesare.fracassi@mccombs.utexas.edu

1 Introduction

Much literature in economics emphasizes specialization and the division of labor as the key drivers of long-run economic growth.[1] The idea is that by letting economic agents increasingly focus on the narrow set of tasks at which they are relatively efficient, aggregate productivity is gradually enhanced. Different strands of the literature have focused on different levels of aggregation: Adam Smith's famous pin-factory example focuses on individual workers;[2] while much international trade literature since David Ricardo focuses on entire countries,[3] building on the seminal concept of comparative advantage.

In our paper we start with a premise that follows the above economic tradition closely. We model an economy that is populated by *business units*, which are taken to be the elementary agent of production, and where each business unit has a specific technology. The more business units are allowed to focus on the economic activities for which their technology is best suited, the higher the output they generate. We refer to this business-unit level of focus as *technological specialization*. It is within this setting that we consider the problem of corporate diversification, and in our model it is possible to generate efficiency gains from joining two separate business units under the same firm. In particular, the combination of complementary technologies enables the firm to efficiently reallocate resources across units, as a response to an uncertain business environment. An implicit assumption of our model is that such reallocation is feasible within firms but not across firms, for example because of greater adverse selection.[4] Joining two business units within the same firm has benefits but also entails organizational-complexity costs. In equilibrium costs and benefits are traded off, allowing our model to make general predictions about patterns of corporate-diversification activity, and in particular understand how it is affected by technological specialization. In a calibration exercise we are able to fit U.S. data relatively well, suggesting the model is a

[1]For an extensive review on this topic, see Yang and Ng (1998).

[2]Smith (1776).

[3]See Ricardo (1817) and Dixit and Norman (1980).

[4]This assumption is in line with an interpretation of the boundaries of the firm as information boundaries, as suggested, for example, in Chou (2007).

useful framework for understanding corporate diversification.

Our setup is a continuous-time matching model of diversifying mergers, where single-segment firms—i.e., firms that only comprise one business unit—are paired up at random according to a Poisson process with exogenous intensity. This approach is in line with search-and-matching models of unemployment (Diamond, 1993; Mortensen and Pissarides, 1994), and has been used before as a model of merger activity (Rhodes-Kropf and Robinson, 2008). Also, and here we once more follow the search-and-matching literature on unemployment, we assume diversified firms refocus at some future random moment in time, where each division becomes again a separate corporation. Whereas the refocusing event is random, it does create value on average, since we also assume that organizational-complexity costs increase over time. This could be interpreted as growing agency costs inside conglomerates, as top managers become entrenched, an assumption in the spirit of papers on the "dark side" of internal capital markets.[5] For simplicity, corporate diversification and refocusing in our model are entirely driven by mergers and spin-offs. Focusing on corporate-restructuring mechanisms is consistent with previous literature: In Graham, Lemmon, and Wolf (2002), almost two thirds of the firms that increase the number of segments implement this strategy via acquisition; and many diversifying mergers are later divested (Ravenscraft and Scherer, 1987; Kaplan and Weisbach, 1992; Campa and Kedia, 2002; Basu, 2010).

One of the distinctive features of our approach is the spatial nature of the model. In particular, the economy is populated by a continuum of business units that are uniformly located on a circle. This location represents the technology type each business unit is endowed with. Business units pursue projects, which are also characterized by a location on the circle, and project output decreases in the distance between the technology and the project (type). This captures the notion that each business unit has a comparative advantage for certain projects, but not for others. The exogenous level of technological specialization determines how narrow the potential range of projects is, and therefore average project output increases

[5] See for example Scharfstein and Stein (2000), Scharfstein, Gertner, and Powers (2002), and Rajan, Servaes, and Zingales (2000).

with technological specialization. Our notion of specialization aims to represent the extent to which market institutions are developed in the economy, such that certain tasks can be "outsourced", leaving the business unit free to focus on the economic activities at which it is relatively efficient. In the model, diversifying mergers generate synergies because business units within the same firm can trade projects whenever this is efficient; this in-house project trade represents within-conglomerate resource reallocation. Thus our approach is close to the internal capital markets literature (Stein, 1997; Scharfstein and Stein, 2000), albeit we consider an ability to reallocate resources other than financial capital.

Diversifying synergies initially increase in cross-division technological distance, which we henceforth refer to as *segment distance*. The intuition for this effect is that complementarity is relatively low if two business units are very similar, since trading projects can only generate limited gains (in fact zero as technologies fully overlap). The initial increase of synergies in segment distance is reduced gradually, until an interior optimum is attained, and afterwards synergies decrease. The intuition for the decrease is in the spirit of diversification literature emphasizing *relatedness* as a source of conglomerate value, starting with Berger and Ofek (1995): In our model project type is drawn from a distribution centered at the business-unit's technological location, which means that very distant technology pairs never exhibit efficiency gains from in-house project trade.

The final aspect of the spatial representation of our approach is the distribution of merger matches. In particular, we assume that merger opportunities only occur within a relatively small neighborhood of the business unit's technology. This "home bias" could be rationalized if adverse-selection concerns grow in technological distance, the explicit modeling of which is outside the scope of our paper. Given a strong enough home bias in merger activity, the model then makes a somewhat counter-intuitive prediction, namely that *observed* diversification synergies increase with segment distance. This is in the spirit of much literature on social and economic networks, where agents who span disconnected environments—"brokers"—obtain significative rents therefrom.[6] Moreover, our empirical analysis is indeed consistent

[6] For a review of these topics, see Burt (2005) and Jackson (2008)

with this prediction, as we detail below.

We now turn to mapping our theoretical setup to data. Our main challenge is to find the appropriate measure for segment distance. We employ the approach in Anjos and Fracassi (2013), who use input-output flows to construct an industry-network representation of the U.S. economy. Segment distance is thus defined as the length of the shortest path connecting a pair of industries in the aforementioned network. There are three main advantages associated with this measure of segment distance. First, it is defined for all industries in the economy, and not just the subset of manufacturing industries. This is important for our purpose of characterizing economy-wide corporate-diversification activity, and so we would not want to employ a technological similarity measure that is only defined for manufacturing, as for example in Bena and Li (2013). Second, our concept of "technology" is quite broad, as in standard macroeconomic models. More specifically, it includes a firm's managerial/organizational technology, which is potentially similar for industries that are close-by in the economy-wide supply chain: For example, suppose two vertically-disconnected industries A and B share a key supplier industry C; then it seems reasonable that a management team of company A would be relatively efficient in managing firm B. Finally, our segment-distance variable also has the advantage of not being overly dependent on the specific industry-classification scheme, unlike the one proposed by Berger and Ofek (1995). In particular, if two industries are focusing on a similar economic activity, one would expect, everything else constant, that these two industries have a similar set of customer and supplier industries. Sharing these indirect connections yields a low segment distance, which thus is capturing how equivalent two industries are in the economy-wide supply chain. Moreover, segment distance generalizes this notion of technological equivalence by also including higher-order indirect connections—customers of customers, customers of suppliers, and so on.

Using the segment-distance variable, we then investigate whether the model provides an accurate description of corporate diversification patterns in the U.S. First we run OLS regressions of conglomerate value on segment distance, and we find a positive association, even after controlling for a host of other factors, including the level of direct vertical relatedness (or intensity of direct linkages). This positive association is consistent with our model as long as the home-bias in merger matches is strong enough. This finding stands in contrast with the usual stance in finance research about *relatedness* (broadly defined), which is usually understood to be a positive factor behind synergies (Berger and Ofek, 1995; Fan and Lang, 2000; Hoberg and Phillips, 2010; Bena and Li, 2013). However, a positive association between relatedness and value is potentially identified by unrelated deals that are motivated, for example, by managerial empire-building; and not all empirical measures of similarity/relatedness necessarily pick up such agency effects to the same extent. Therefore, our view and the positive-relatedness view are not necessarily inconsistent or mutually exclusive, albeit this argument does suggest that more research is required to understand what exactly each empirical measure is picking up.

Second, we calibrate our model using U.S. data on corporate diversification for the period 1990-2011. We start our data in 1990 because we require NAICS classification codes in order to construct the segment-distance variable. The calibration is able to match important magnitudes that characterize aggregate corporate-diversification activity: the proportion of assets allocated to single-segment firms in the economy, the so-called "diversification discount", and the above-mentioned cross-sectional association between segment distance and value. The calibration achieves this matching using a standard level for the discount rate, a reasonable frequency of merger activity for the representative firm, and a reasonable average level for Tobin's Q. Also, although we match the diversification discount, this is somewhat of an apparent discount, since all firms are rational value-maximizers. The discount stems from the fact that organizational-complexity costs grow over time and thus are heavily discounted when firms are making decisions about diversification. Nothwithstanding the rationality of this decision for shareholders, the average conglomerate becomes inefficient over time, and due to (ex-post) managerial entrenchment it commands a relatively low value. But ex-post managerial entrenchment and its costs are indeed taken into account by shareholder-aligned single-segment firms. This rational explanation for the diversification discount is in the spirit of Anjos (2010).[7]

The calibrated model has relatively sharp implications for the importance of corporate diversification as an economic activity. In particular, the option to diversify accounts for slightly more than 4% of single-segment firm value. As an upper bound, we find that the best-possible matches occurring in the model lead to shareholder value creation of about 7% (this would be the predicted merger announcement return).

In addition to trying to understand the levels of corporate-diversification variables, we also investigate the trends in these variables. Our evidence strongly suggests that the proportion of assets allocated to single-segment firms has been increasing over the period 1990-2011; and, also, that segment distance for a representative conglomerate has been slowly decreasing. In addition, data also suggests, albeit with more noise, that the diversification discount has been decreasing, and that the Tobin's Q of the representative firm in the economy has become higher. A comparative-statics exercise using our calibrated model sheds some light on these trends. More specifically, an increase in technological specialization—the key parameter in our model—generates the aforementioned patterns. Moreover, such patterns crucially depend on the existence of a discount, or, equivalently, on high enough marginal costs of organizational complexity. If these costs are shut down, the model actually predicts that the pervasiveness of conglomerates is independent of the level of technological specialization, depending exclusively on match- and break-up rates, as well as the discount rate.

The remainder of the paper is organized as follows. Section 2 develops the theoretical setup, which entails a model for the relationship between segment distance and flow synergies; and a model for the process through which diversification activity occurs and firm boundaries change. Section 3 documents empirical evidence and performs a calibration. Section 4 analyzes empirical trends and attempts to rationalize them with a simple comparative-statics exercise. Section 5 concludes. An appendix contains all proofs and summary statistics.

[7]Other papers have proposed rational explanations for the discount using dynamic models; see for example Matsusaka (2001), Bernardo and Chowdhry (2002), Maksimovic and Phillips (2002), and Gomes and Livdan (2004).

2 Model

This section contains our theoretical setup. We start by developing a static equilibrium model for flow payoffs (section 2.1), which we then embed in a dynamic search-and-matching framework (section 2.2). The solution to the dynamic model is presented in section 2.3.

2.1 Flow payoffs

The economy comprises a continuum of business units (henceforth BUs), where BU i is characterized by a location α_i on a circle with measure 1.[8] The different locations on the circle represent different technologies, which enable firms to pursue profitable project opportunities. Our notion of technology is broad, as in standard macroeconomic models, and we define it more concretely in the empirical section.

Business units are organized either as a single-BU firm or as a two-BU (or two-segment) corporation, which we term a conglomerate. We take the organizational forms as given for now; these are endogenized later (section 2.2). The next two subsections further characterize the static payoffs of single-segment and diversified firms.

2.1.1 Single-segment firms

Each BU in the economy undertakes one project,[9] and this project is also characterized by a location in the technology circle, denoted by α_{P_i}. This location represents the *ideal* technology, that is, the technology that maximizes the project's output. The location of the project is drawn from a uniform distribution with support $[\alpha_i - \sigma, \alpha_i + \sigma]$, and the distribution being centered at α_i implies that on average BUs are well-equipped to implement the projects they are presented with. The higher σ is, the higher the uncertainty about the ideal technology required by projects, and we interpret the inverse of σ as the degree of

[8] 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.

[9] An implicit assumption of our model is 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).

technological specialization, which in our model is taken as exogenous. Specialization thus refers to how much productive units are able to focus on those activities (projects) for which their technology is best suited. For tractability we assume $\sigma < 1/4$, which greatly simplifies the analysis.[10] The support of the distribution for project location corresponds to the dashed arc in figure 1.

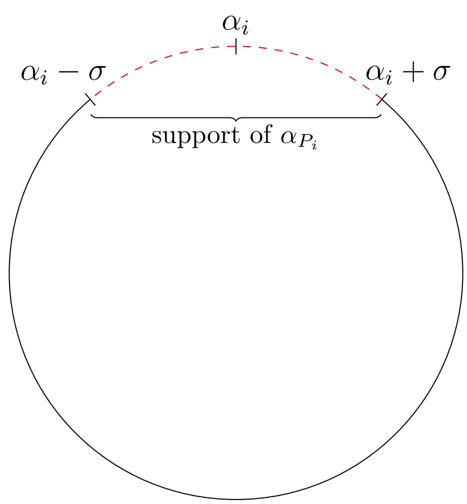

Figure 1: Technologies and Projects: Spatial Representation. The figure depicts a circle where both projects and business units are located. The location of the business unit represents its technology, whereas the location of projects represents the ideal technology to undertake that particular project. The figure also shows that business units draw projects from locations close to their technology.

If BU i is organized as a single-segment firm, then its profit function is given by the following expression:

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

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

[10]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.

the expected profits of a single-BU firm, denoted as π_0, are given by

$$\pi_0 := \mathrm{E}\left[\pi_i\right] = 1 - \phi\frac{\sigma}{2}. \tag{2}$$

Equation (2) shows that an increase in specialization (decrease in σ) leads to higher profits, which attain their maximal level of 1 with "full specialization" ($\sigma = 0$).

As will become gradually apparent, σ is the key parameter in our model. The fact that we focus on specialization, albeit with a novel formal approach, is in the spirit of much of economic tradition (see introduction for discussion).

2.1.2 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 single-segment 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, which we detail in section 2.2. 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 (i.e., before HQ costs), taking *segment distance*—in the technology circle—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 $z \le \sigma$,

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

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

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

3. If $z > 2\sigma$,

$$\pi_1^G(z) = \pi_0. \tag{5}$$

Figure 2 depicts the relationship between segment distance and average division profits, and illustrates the natural ambiguity in this relationship. If distance is too low, there are many efficient project transfers, however the average gain of each transfer is small. If distance is too high, then realized project transfers correspond on average to a large gain; however, each division is usually the closest to the projects it generates, and so transfers are rare. The optimum distance trades off the frequency of desirable transfers with the average gain of each transfer. Proposition 2 shows that optimal (static) segment distance is a simple proportion of project-type uncertainty σ, which is intuitive.

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

$$z^* = \sigma\left(2 - \sqrt{2}\right), \tag{6}$$

with associated expected BU profit of

$$\pi_1^G(z^*) = 1 - \phi\sigma\frac{(4-\sqrt{2})}{6}. \tag{7}$$

If σ is interpreted as a measure of the inverse of specialization, then an increase in specialization (lower σ) 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. However, this does not necessarily imply that one should observe fewer conglomerates, as we show later in the solution to the dynamic problem.

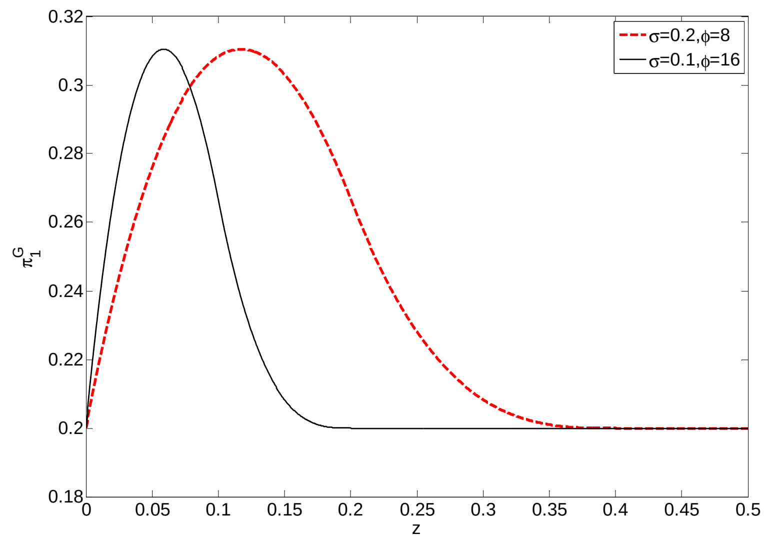

Figure 2: Segment Distance and (Static) Profits. The figure plots profits π_1^G as a function of segment distance z.

It is not clear which empirical relationship between segment distance and profits is implied by this simple static model. Inspecting figure 2, 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—say because managers pursue zero-synergy mergers for empire-building motives—then actually the average relationship between segment distance and value would be negative. 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 2 one observes that the relationship between segment distance and profits is *relative* to σ. As long as the product $\phi\sigma$ is constant, the maximal value of synergies is the same (see proposition 2). Therefore, holding the product $\phi\sigma$ constant, it would not be possible to distinguish between an economy where σ is high and the distribution of firms has wide support (dashed curve of figure 2) from an economy with low σ but where the distribution of firms has narrow support (solid curve of figure 2). This point is important for our calibration, where given the argument just outlined we set σ at an arbitrary level.

2.2 Dynamics

2.2.1 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 technologies, 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 assumption that all firms have the same scale and that adjustment costs are infinite. While 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 presented in section 3.2. 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 single-BU firms that at some point in the past found it optimal to merge. 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). Each pair of extant single-segment firms may be presented with a potential merger opportunity according to a Poisson process with intensity λ_0. If a meeting between two single-segment firms occurs, a merger happens as long as it creates value, and surplus is shared equally across merging partners. For simplicity, and following search-and-matching models of employment, mergers are reversed according to an exogenous Poisson process with intensity λ_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 single-segment 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 uniform distribution with support $[0, z_{max}]$, where $z_{max} = \eta\sigma$. Our assumption implies that the set of potential partners is drawn from a location that is a neighborhood of the BU's business environment (thus a function of σ), and in general we consider $z_{max} << 1/2$, which implies a "home bias" in diversification activity. This assumption aims to capture the notion that managers are more "confortable" engaging in merger deals within a neighborhood of their business environment. This home bias could be rationalized in a model with heterogeneous firm quality and where adverse-selection concerns increase with technological distance, the explicit modeling of which is beyond the scope of our paper.

2.2.2 The cost of headquarters

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

$$\pi_1^G(z_\tau) - \beta_0 e^{\beta_1(t-\tau)} \quad , \forall t \in [\tau, \tau^+], \tag{8}$$

where π_1^G is the equilibrium profit function from the static setup, i.e. equations (3)-(5), and the second term corresponds to HQ costs. According to the expression above, HQ starts off with some cost rate β_0, but this cost rate increases over time, at rate β_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 later costs are heavily discounted when making diversification choices, however they show up in *unconditional* average conglomerate value. For simplicity we assume there are no HQ costs for single-segment 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.

A critical implicit assumption in our modeling of HQ costs is that they are independent of segment distance. This is consistent with the standard notion of decreasing returns to scale, and also with the findings in Sanzhar (2006), who shows that much of the inefficiencies associated with conglomerates are driven by the fact that they are multi-unit corporations—and not specifically because they combine divisions from different industries or geographies.

2.3 Solving the model

First let us state the individual optimization problem. Since (matched) firms share merger surplus equally, the optimization problem from the perspective of business unit i is as follows:

$$\begin{aligned} J_t &= \sup_{\{\tau\}} \left\{ \mathrm{E}_t \left[\int_{u\in[t,+\infty]\cap\{[\tau,\tau^+]\}} e^{-r(u-t)} \left[\pi_1^G \left(z_{\sup\{\tau<u\}} \right) - \beta_0 e^{\beta_1 (u-\sup\{\tau<u\})} \right] \, du + \right.\right. \\ &\quad \left.\left. + \int_{u\in[t,+\infty]\setminus\{[\tau,\tau^+]\}} e^{-r(u-t)} \pi_0 \, du \right] \right\}, \end{aligned} \tag{9}$$

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, τ^+ returns the first time after τ at which a split takes place, and $z_{\sup\{\tau<t\}}$ is the distance of the two divisions inside the diversified firm.

The solution concept we employ is Markov Perfect Equilibrium (see for example Maskin and Tirole, 2001), which is outlined in definition 1.

Definition 1 *(Equilibrium) A Markov Perfect Equilibrium of this economy is characterized by an unchanging proportion of single-segment 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 (9).*

The next proposition characterizes the equilibrium value functions for single-segment and diversified BUs.

Proposition 3 *In an equilibrium with no mergers, the value of single-segment firms J_0 is equal to π_0/r. In an equilibrium with mergers, the optimal policy of single-segment firms is characterized by accepting matches with segment distance in an interval $[z_L, z_H]$. In such an equilibrium, the time-t value of a business unit inside a diversified firm, J_1, is a function of the segment distance at which the merger took place (z) and the current duration of the merger ($d := t - \tau$):*

$$J_1(z, d) = \frac{\pi_1^G(z) + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0}{r + \lambda_1 - \beta_1} e^{\beta_1 d} \tag{10}$$

The value of single-segment firms J_0 is characterized as

$$J_0 = \frac{1}{r}\left[\frac{r + \lambda_1}{r + \lambda_1 + \lambda_0 q}\pi_0 + \frac{\lambda_0 q}{r + \lambda_1 + \lambda_0 q}\left(\overline{\pi}_1^G - \beta_0 \frac{1}{1 - \frac{\beta_1}{r+\lambda_1}}\right)\right], \tag{11}$$

with q the probability of merger acceptance and $\overline{\pi}_1^G$ the average diversified-BU gross profit rate:

$$q := \frac{z_H - z_L}{z_{max}} \tag{12}$$

$$\overline{\pi}_1^G := \int_{z_L}^{z_H} \frac{1}{z_H - z_L} \pi_1^G(z)\, dz \tag{13}$$

Equation (11) describes the equilibrium value of single-segment firms, which embed the value of the option to diversify 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 single-segment cash-flow rate π_0 and an average "time-adjusted" cash-flow rate of diversified BU's, given by the term

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

The above term increases in the average gross profit of diversified BUs, and decreases in starting HQ costs β_0 and the rate at which these costs grow β_1. In equation (11), what determines the weight of the single-BU cash-flow rate, relative to the diversified cash-flow rate, is how frequent mergers and break-ups are, which is influenced by λ_0, λ_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 single-segment.

Proposition 4 characterizes equilibrium pervasiveness of merger and diversification activity in the economy.

Proposition 4 *The following three results obtain in a Markov Perfect Equilibrium:*

1. *The proportion of single-segment firms in the economy is given by*

$$p = \frac{\lambda_1}{\lambda_1 + \lambda_0 q}. \tag{14}$$

2. *There exists a threshold W, defined as*

$$W := \frac{6\beta_0}{\left(\sqrt{2} - 1\right)\left(1 - \frac{\beta_1}{r+\lambda_1}\right)}, \tag{15}$$

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

3. *For the special case where $\beta_0 = 0$ (no HQ costs), p and q do not depend on either ϕ*

and σ, and the optimal merger policies are proportional to σ:

$$z_L = \theta_L \sigma, z_H = \theta_H \sigma, \tag{16}$$

with $\theta_H \geq \theta_L \geq 0$, and θ_L, θ_H functions of r, λ_0, λ_1, and η.

The first result in proposition 4 shows that the steady-state proportion of single-segment firms is a simple function of match- and break-up rates, as well as merger acceptance probability. The second result in proposition 4 shows that mergers only take place if either the location of projects is highly uncertain (high σ) or the cost of project-technology misfit is high (ϕ). 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 single-BU firms. These benefits of diversification are compared to its costs, gaged by initial HQ cost β_0, and the rate at which this cost grows β_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 λ_1). In this model, knowing ex ante that the conglomerate will last little actually fosters diversification, which follows directly from our assumption that HQ costs grow over time. Finally, the third result in proposition 4 highlights how the equilibrium pervasiveness of conglomerates in the economy does not depend on ϕ and σ whenever there are no HQ costs. This result follows from the fact that single-segment and diversified-firm value both become linear in the product $\phi\sigma$. This linearity is broken when there are HQ costs (which are independent of both ϕ and σ) and carries important implications for the comparison of model with data (see section 4).

Proposition 5 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 λ_1 and the rate at which HQ costs grow β_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 5 *If $\lambda_1 > \beta_1$, the average value of diversified firms in the economy is given by*

$$\mathrm{E}[J_1] = \frac{\overline{\pi}_1^G + \lambda_1 J_0}{r + \lambda_1} - \frac{\beta_0 \lambda_1}{(\lambda_1 - \beta_1)(r + \lambda_1 - \beta_1)}, \tag{17}$$

otherwise $\mathrm{E}[J_1] = -\infty$.

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

Proposition 6 *The equilibrium specified in definition 1 always exists and is unique.*

3 Data and calibration

This section turns to data, where we first describe the construction of our empirical proxy of segment distance (or distance in technologies) and document its association with firm value (section 3.1). Afterwards we proceed to calibrate the model developed in the previous section, and learn what the calibration implies about the economic importance of corporate diversification (section 3.2).

3.1 Data description and variable construction

3.1.1 The segment-distance variable

The model we developed in the previous section takes the following twofold view: (i) the appropriate combination of technologies is important for firm performance; and (ii) merger activity and corporate diversification are important forms of achieving the right combination of technologies. The question now is: How to measure similarity across technologies within a conglomerate? We employ the approach in Anjos and Fracassi (2013), who use input-output flows to construct an industry-network representation of the U.S. economy. With this approach, conglomerate segment distance is defined formally as follows:

$$Seg.Dist. = \frac{\sum_{i \in \mathcal{I}} \sum_{j>i \wedge i \in \mathcal{I}} l_{ij}}{M(M-1)/2}, \tag{18}$$

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 thus indirect linkages matter.[11]

We believe there are three main advantages associated with the above measure of segment distance. First, it is defined for all industries in the economy, and not just the subset of manufacturing industries. This is important for our purpose of characterizing economy-wide corporate-diversification activity, and so we would not want to employ a technological similarity measure that is only defined for manufacturing, as for example in Bena and Li (2013). Second, our concept of "technology" is quite broad, as in standard macroeconomic models. More specifically, it includes a firm's managerial/organizational technology, which is potentially similar for industries that are close-by in the economy-wide supply chain: For example, suppose two vertically-disconnected industries A and B share a key supplier industry C; then it seems reasonable that a management team of company A would be relatively efficient in managing firm B. Finally, our segment-distance variable also has the advantage of not being overly dependent on the specific industry-classification scheme, unlike the one proposed by Berger and Ofek (1995). In particular, if two industries are focusing on a similar economic activity, one would expect, everything else constant, that these two industries have a similar set of customer and supplier industries. Sharing these indirect connections yields a low segment distance, which thus is capturing how equivalent two industries are in the economy-wide supply chain. Moreover, segment distance generalizes this notion of technological equivalence by also including higher-order indirect connections—customers of

[11] The reader is referred to Anjos and Fracassi (2013) for further details on the empirical implementation of segment distance. In particular, our segment-distance measure employs normalized inter-industry flows, which we believe are the most appropriate in our setting. We use the 1997 input-output tables for flow construction (as do Ahern and Harford, 2012 and Anjos and Fracassi, 2013).

customers, customers of suppliers, and so on.

Finally, with a networks approach to distance one can also distinguish the effect of technological proximity elaborated above from more-standard arguments for vertical integration, which we argue are more appropriately proxied for by the intensity of the direct connection.[12]

3.1.2 Segment distance and conglomerate value

In this section we document the relationship between segment distance and firm value, using U.S. data from 1990 to 2011. We start our data in 1990 because we require NAICS classification codes in order to construct the input-output-based industry network. Our key independent variable, segment distance, is computed using the detailed input-output tables for the year 1997. 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 single-segment firms; the idea behind this variable is to control for industry-specific valuation patterns. Summary statistics are presented in table A.1 in the appendix.

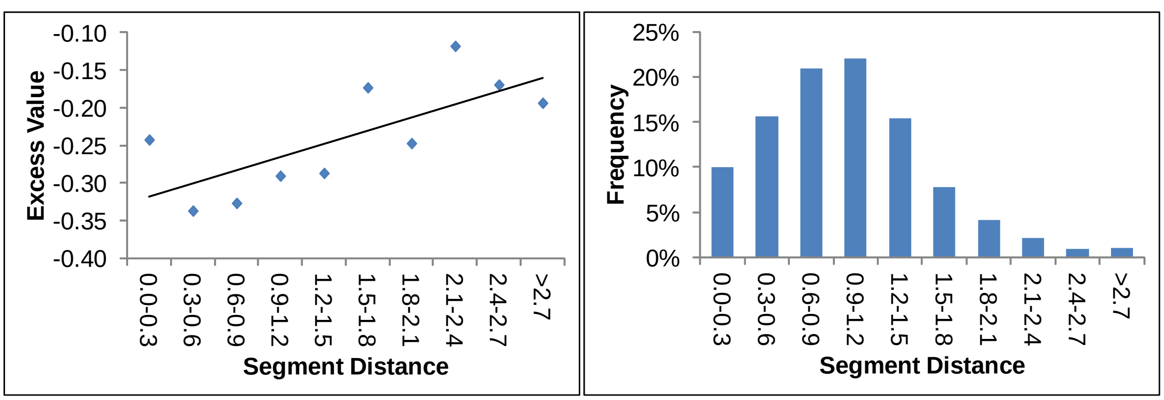

Figure 3: Segment Distance and Excess Value. The left panel shows conglomerate average excess value, conditional on segment-distance class. The right panel shows segment-distance distribution. Excess value is defined as the log-difference between the Tobin's Q of a conglomerate and the Tobin's Q of a similar portfolio of single-segment firms, following Berger and Ofek (1995). Segment Distance is the average input-output-based distance across conglomerate segments, following Anjos and Fracassi (2013).

The left panel in figure 3 shows the empirical association between segment distance and

[12]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.

conglomerate valuation, and the right panel describes the segment-distance distribution. 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, more importantly, single-segment firms are perfectly aligned with shareholders. The relationship between segment distance and firm value is economically significant: Sorting by segment distance, the difference in excess value between above-median and below-median conglomerates is about 6%. Excess value is approximately equal to

$$\frac{Q_{conglomerate} - Q_{benchmark}}{Q_{benchmark}},$$

and on average the Tobin's Q of conglomerates is 28% lower than that of the (single-segment) benchmark. Therefore, a difference of 6% in excess value is about $6\%/0.72 \approx 8\%$ in value for the average conglomerate. This positive association between segment distance and excess value is consistent, in light of our model, with merger opportunities taking place in a relatively close neighborhood of the firm's core activities. There are plausible reasons for why this "home bias" would take place, for example adverse selection being more of a concern for distant mergers. Also, the initially positive association between segment distance and frequency, shown in the right panel of figure 3, is consistent with the notion that firms prefer intermediate-distance combinations to low-distance combinations. That the frequency afterwards decreases is however not necessarily a function of firms not preferring high-distance deals, *per se*. In particular, it seems reasonable that fewer M&A deals are free from serious adverse-selection issues as distance increases (explaining the low frequency); but, for those where adverse selection is indeed not a concern, then one observes relatively high synergies (explaining high Tobin's Q for high-segment-distance firms).

The simplistic analysis in the left panel of figure 3 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.[13] For ease of interpretation, all variables have been standardized.

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 single-segment 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)
Segment Distance	0.043***	0.037**	0.035*	0.084***
	(2.80)	(2.07)	(1.95)	(3.53)
N. Segments		-0.054***	-0.063***	-0.080***
		(-3.10)	(-3.42)	(-3.90)
Related Segments		0.051***	0.043**	0.016
		(2.74)	(2.38)	(0.82)
Vert. Relatedness		0.023*	0.017	0.073***
		(1.90)	(1.34)	(3.05)
Excess Centrality		0.040**	0.036*	0.062**
		(2.13)	(1.90)	(2.27)
Excess Assets			0.055***	-0.059
			(2.78)	(-1.54)
Excess EBIT/Sales			-0.091***	-0.026***
			(-9.18)	(-2.70)
Excess Capex/Sales			0.015***	0.029***
			(3.02)	(8.73)
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	No	Yes
R^2	0.015	0.018	0.030	0.028
Obs.	22,425	22,425	21,516	21,516

Specification (1) in table 1 presents the correlation between segment distance and excess value, but now controlling for year fixed effects. Specification (2) 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

[13]Results are economically similar if we restrict the analysis to 2-segment conglomerates only, which are closer to our model.

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 vertical relatedness loads only on the intensity of *direct* bilateral links. Model (2) also includes the excess centrality measure in Anjos and Fracassi (2013), which aims to capture a conglomerate's informational advantage relative to single-segment firms. The coefficient of segment distance remains statistically and economically significant after including year fixed effects and other diversification characteristics. According to model (2), a one-standard-deviation increase in segment distance is associated with an increase of about 0.037 standard deviations in excess value. Excess value has a standard deviation of 0.66, so this corresponds to an increase of about 0.024 in excess value, which is about $0.024/0.72 \approx 3.3\%$ in firm value for the average conglomerate.

Specification (3) adds financial variables to the regression, constructed according to the the approach recommended in Gormley and Matsa (2013),[14] and specification (4) includes firm fixed effects, which 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.[15] The most stringent regression is also the one where segment distance has the strongest economic effect, which more than doubles relative to model (2). Also, in model (4) segment distance is the independent variable with the highest absolute coefficient, i.e., it is the variable from the set of regressors that most explains variation in excess value.

In table 1 the coefficients on number of segments, related segments, and vertical relatedness are all consistent with previous literature: relatedness is associated with higher firm value. This begs the question of why the results are qualitatively different with segment distance and excess centrality. Our theory notwithstanding, it is certainly plausible that

[14]Results are however similar if we use raw financial conglomerate variables, instead of computing excess measures.

[15]With the caveat that time-varying managerial skills or firm organizational capital could still render our results spurious.

firms engaging in totally disconnected (i.e., zero-synergy) business combinations do so for the wrong reasons, e.g., managerial empire-building. Everything else constant, this implies a positive association between relatedness and value. However, we also believe that it is plausible that highly-related business combinations are redundant and should display low complementarity and therefore low value. More importantly, the co-existence of the two arguments suggest that it is possible for some measures of relatedness/similarity to pick up mostly agency problems, whereas others would pick up mostly the benefits of combining complementary technologies (segment distance) or non-redundant information (excess centrality).

3.2 Calibration

The data we use for the baseline calibration is the same used in section 3.1. Summary statistics are presented in the appendix. There are a total of eight parameters to calibrate: r, η, λ_0, λ_1, β_0, β_1, ϕ, and σ. 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.

As explained in section 2.1, it would be hard to separately identify σ, which determines the location of positive synergies in the z space, from the location of the distribution of matches;[16] so we opt to do a calibration where everything is *relative* to σ, which we arbitrarily set at 0.2. Finally, we use the value of single-BU firms to pin down ϕ. In data, the value of single-segment firms (per unit of capital) has an average of 2.6. 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 ϕ, which we set at 8; this yields a lower bound for the value of

[16] See figure 2 and related text.

single-segment 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

$$\frac{0.1 \times 2}{0.1 - 0.02} = 2.5, \tag{19}$$

which is close to 2.6, its data counterpart.

We are left with four parameters to calibrate: η, λ_0, β_0, and β_1. We use four moments from data in order to identify these parameters: (i) the proportion of single-segment firms in the economy, which we measure using the in-sample average proportion of book assets owned by single-segment corporations, approximately 55%; (ii) the average excess value in the economy, which in our sample is -0.28; (iii) the difference in excess value between the top-50% conglomerates in terms of segment distance, relative to the bottom 50%, which in our sample is 0.06 (let us denote this moment by "Δ Excess Value"); and (iv) the likelihood that a firm is involved in a takeover, which for example in Edmans, Goldstein, and Jiang (2012) is 6% (per year). The model counterparts to (i)-(iii) are straightforwardly computed using results from section 2.3. As for the probability that a firm engages in at least one merger over the course of a year, in the model this corresponds to 1 minus the probability that the firm does not engage in any merger, which is given by

$$\sum_{k=0}^{\infty} \Pr\{matches = k\}(1-q)^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k(1-q)^k}{k!} = \frac{e^{-\lambda}}{e^{-\lambda(1-q)}} \underbrace{\sum_{k=0}^{\infty} \frac{e^{-\lambda(1-q)}[\lambda(1-q)]^k}{k!}}_{=1} = e^{-q\lambda}. \tag{20}$$

Table 2 summarizes the choice of parameters, and table 3 reports key moments.

The calibration delivers a value for single-BU firms of $J_0 = 2.09$, which implies that the value of the option to diversify is about 4.3%, since the value of single-BU firms without this

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

Parameter	Value
r	0.10
η	0.60
λ_0	0.11
λ_1	0.10
β_0	0.04
β_1	0.07
ϕ	8.00
σ	0.20

Table 3: Model outputs and data. The table shows key moments, both in the calibration and in data. "Single-Seg. Value" is the Tobin's Q of single-segment firms; "Prop. Single-Seg." is the proportion of assets in the economy allocated to single-segment firms; "Av. Excess Value" is the unconditional excess value of conglomerates; "Δ Excess Value" is the difference in excess value between above-median-segment-distance and below-median-segment-distance conglomerates; and "Probab. of M&A" stands for the likelihood that a single-segment BU engaged in at least one merger deal.

Moment	Model Counterpart	Calibration Output	Data/target
Single-Seg. Value	J_0	2.09	2.00
Prop. Single-Seg.	p	59%	55%
Av. Excess Value	$\frac{E[J_1]-J_0}{J_0}$	-0.28	-0.28
Δ Excess Value	$\frac{E[J_1\|z>z_{median}]-E[J_1\|z\leq z_{median}]}{J_0}$	0.05	0.06
Probab. of M&A	$1-e^{-\lambda q}$	6.6%	6.0%

option is simply π_0/r, in our case 2. In order to match data, we use a z_{max} that is quite low (recall: $z_{max} = \eta\sigma$), actually around the static optimal segment distance (for our choice of ϕ and σ). z_{max} needs to be low such that high-segment-distance conglomerates display on average a higher valuation than low-segment-distance conglomerates; if for example we set $z_{max} = 2\sigma$, this relationship becomes essentially flat on average (see figure 2). Thus in our calibration, a merger opportunity at a high distance is good news. In fact, and using equation (10), the value of a *brand new* conglomerate with $z = z_{max}$ is

$$\frac{\pi_1^G(z_{max}) + \lambda_1 J_0}{r+\lambda_1} - \frac{\beta_0}{r+\lambda_1-\beta_1} \approx 2.24,$$

an increase of about 7% relative to single-segment firm value J_0. 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 single-segment 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 corporate diversification, frictions notwithstanding, is still an important and valuable economic activity.

4 Corporate-diversification trends and specialization

In the previous sections we developed and calibrated a model of corporate diversification, with the implicit assumption of time invariance. This section explores how corporate diversification evolves over time (section 4.1), where we attempt to characterize empirical trends for key corporate-diversification aggregates, and use the model to try to understand these patterns (4.2). Since our model is indeed time-invariant, the way we explore it is by means of a comparative-statics exercise, where we focus on how an increase in specialization (i.e., decrease in σ) changes equilibrium outcomes. We focus on specialization not because we believe the other primitives of the model are not subject to change over time, but because specialization (or project-technology-fit uncertainty) is indeed the key driver of our model; and thus it seems logic to focus on this dimension. Finally we want to add a caveat to the subsequent analysis, namely that for a proper quantitative exercise we would need to incorporate time variation in σ explicitly into the model. This caveat notwithstanding, we do not believe this is crucial for the type of analysis we perform, which is essentially qualitative.

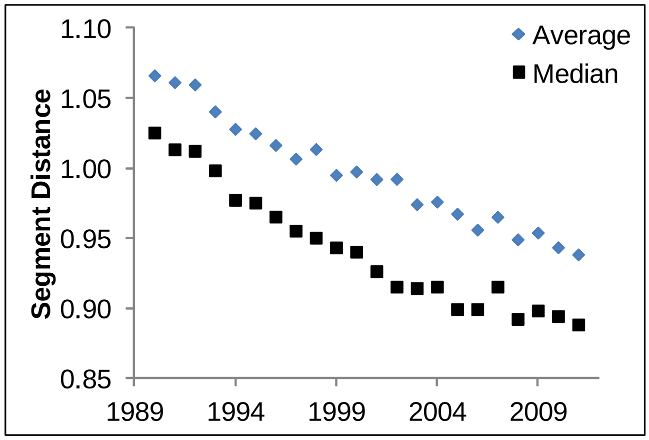

Figure 4: Evolution of Segment Distance. The figure shows average and median segment distance, for the period 1990-2011. Segment Distance is the average input-output-based distance across conglomerate segments, following Anjos and Fracassi (2013).

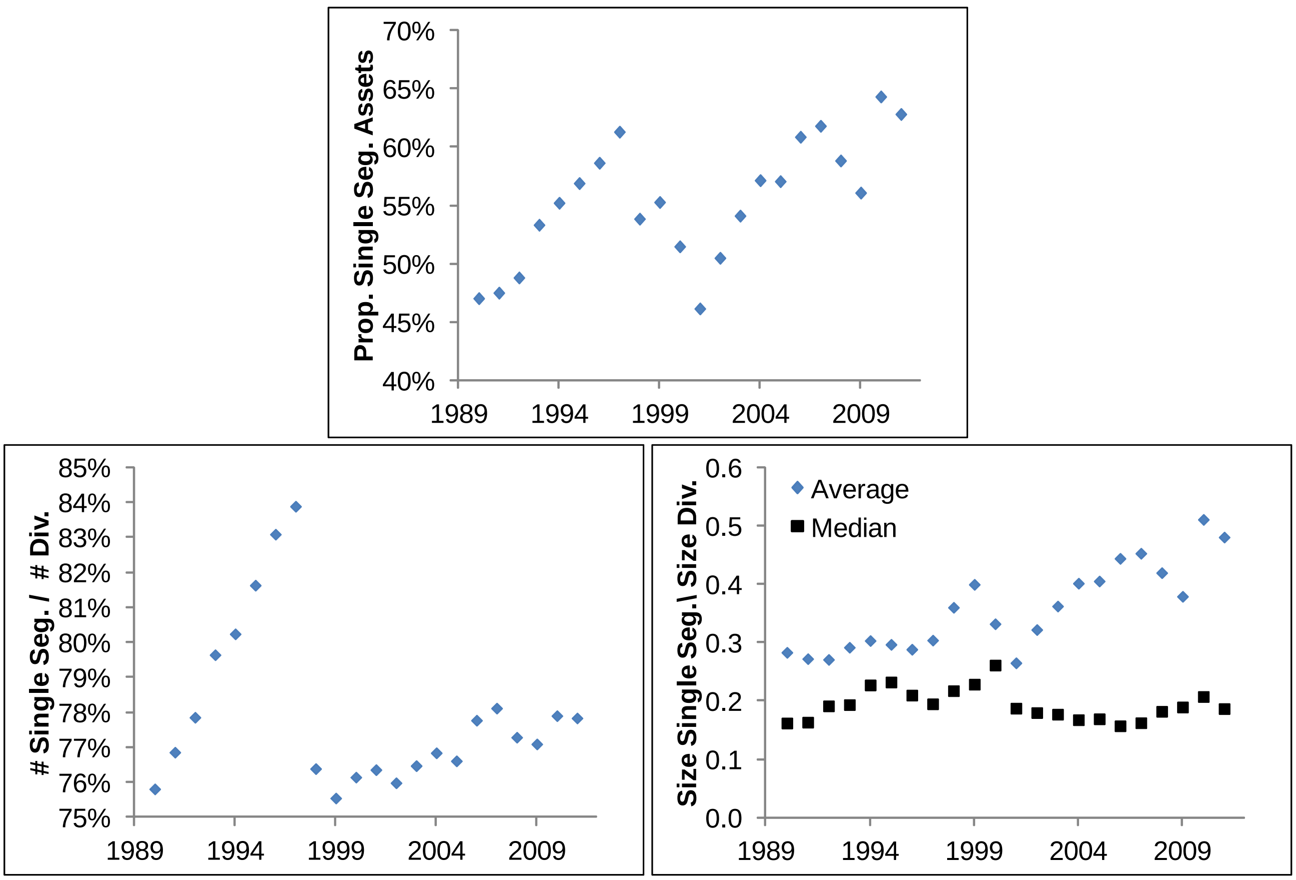

Figure 5: Pervasiveness of Single-Segment Firms. The top panel shows the proportion of total assets in the economy allocated to single-segment firms, for the period 1990-2011. The bottom-left panel shows the ratio the fraction of firms that are single-segment. The bottom-right panel shows the size ratio between single-segment and diversified firms, using the average and the median.

4.1 Empirical time-series patterns

We analyze three aggregate dimensions of corporate diversification: segment distance (figure 4), proportion of single-segment firms (figure 5), and valuation (figure 6). The empirical

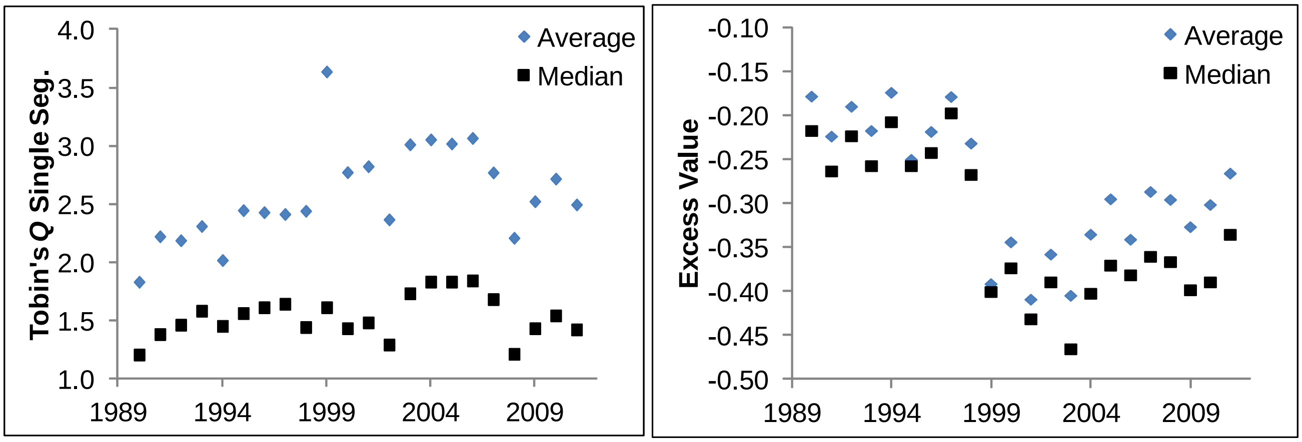

Figure 6: Evolution of Valuation Measures. The left panel shows the Tobin's Q of single-segment firms for the period 1990-2011 (average and median). The right panel shows conglomerate excess value, which is defined as the log-difference between the Tobin's Q of a conglomerate and the Tobin's Q of a similar portfolio of single-segment firms, following Berger and Ofek (1995).

evolution of segment distance is quite uncontroversial and intuitive: Figure 4 shows that for the period 1990-2011 there was a gradual, almost linear decrease in this variable. The trend is the same irrespective of whether we look at averages or medians. We also note that the decrease is small, at least compared with the cross-sectional dispersion of this variable: Segment distance for a representative conglomerate dropped slightly over 10% over a 21-year period, whereas the unconditional standard deviation of segment distance (i.e., including cross-sectional variation) is more than 55% of the unconditional mean. We believe that this slow gradual decline is consistent with a view that specialization is slowly but steadily increasing in the economy. The view that economic growth is mainly a product of gains in specialization is espoused by several economists (see introduction for discussion).

Next we turn to the pervasiveness of corporate diversification activity, which is shown in figure 5. The top panel shows the evolution of the variable we used in the calibration, namely the proportion of book assets allocated to single-segment companies. Whereas a positive trend is present, the data is noisy and apparently cyclical. This is not so much a feature of underlying economic forces, but rather mostly a consequence of the change in segment-reporting requirements introduced in 1997.[17] The bottom-left panel of the same figure plots the fraction of firms classified as single-segment. There is a clear discontinuity

[17] From SFAS 14 to SFAS 131 (see Sanzhar (2006) for more details about the rule changes).

in 1998, consistent with the change in reporting requirements. For each subperiod, the bottom-left panel shows a clear positive trend, albeit the trend is suspiciously strong for early years. This may be related to an attempt by some conglomerates to try to appear as single-segments, in line with Sanzhar (2006). The bottom-right panel of figure 5 plots the asset-size ratio of single-segment to diversified corporations, using the average and the median. For the ratio using medians there is little variation over time; whereas the ratio using averages exhibits a clear upward trend. In summary, we believe the evidence indicates a generalized increased of single-segment activity in the economy.

Finally we analyze valuation trends for both single-segment and diversified firms. The left panel of figure 6 plots Tobin's Q for single-segment firms, both average and median. Our sample excludes extreme outliers that would bias the excess-value analysis, but we note that the inclusion of such outliers would only make the apparent trend in averages even stronger, namely an increase in Tobin's Q. The plot using medians still shows a slightly positive trend, although the effect is weaker. The right panel of figure 6 plots excess value, that is, the log-difference between the value of the conglomerate and the value of a comparable portfolio of single-segment firms. This ratio exhibits a strong discontinuity around the introduction of the new segment-reporting requirements. In the first sub-period there is no apparent trend in excess value, which could potentially be explained by the fact that many single-segment firms are actually misclassified conglomerates. Inclusion of conglomerates in the single-segment sample could make the excess-value variable very noisy (and potentially biased), obscuring any eventual trend. The second subperiod shows a clear upward trend in excess value, both using medians and averages.

We finish this analysis by summarizing our findings. There is strong evidence that, over time and for the period 1990-2011, (i) diversified firms tend to exhibit lower segment distance; and (ii) the proportion of assets allocated to single-segment firms is increasing. The evidence also suggests that both single-segment Q and conglomerate excess value are rising over time.

4.2 Model: comparative statics on σ

This section investigates the implications of varying σ around the calibration choice, and determines whether it can qualitatively account for the trends in key corporate-diversification variables. The main results are presented in figure 7, whereas figure 8 shows how the results would change if there were no conglomerate-specific inefficiencies (HQ costs in the model).

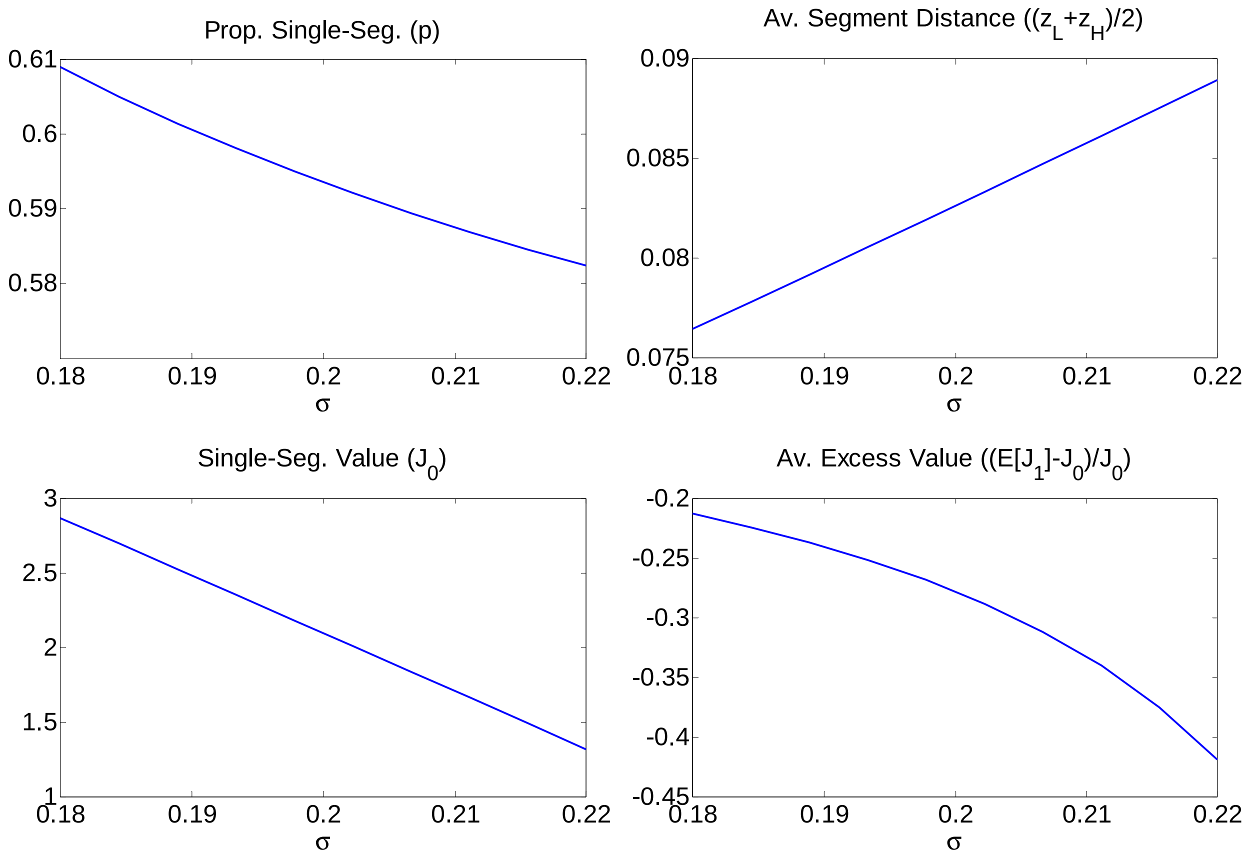

Figure 7: Varying σ: Model Outputs (1/2). The figure shows four key equilibrium outcomes of the model, for σ in an interval of $\pm 10\%$ around the main calibration choice; all other parameters are kept at their original levels (table 2). The top-left panel shows the proportion of single-segment assets in the economy; the top-right panel shows the average diversified-firm segment distance; the bottom-left panel plots the value of single-segment firms; and the bottom-right panel plots conglomerate excess value.

The top-left panel of figure 7 shows that a decrease in σ, which we interpret as an increase in specialization, leads to a higher proportion of single-segment firms. This is in line with the trend in data, and the intuition for the result is straightforward: as σ reduces, the benefits of combining dissimilar technologies are lower relative to HQ costs, and thus in equilibrium one observes fewer conglomerates. Indeed, if we go back to one of the analytical results in

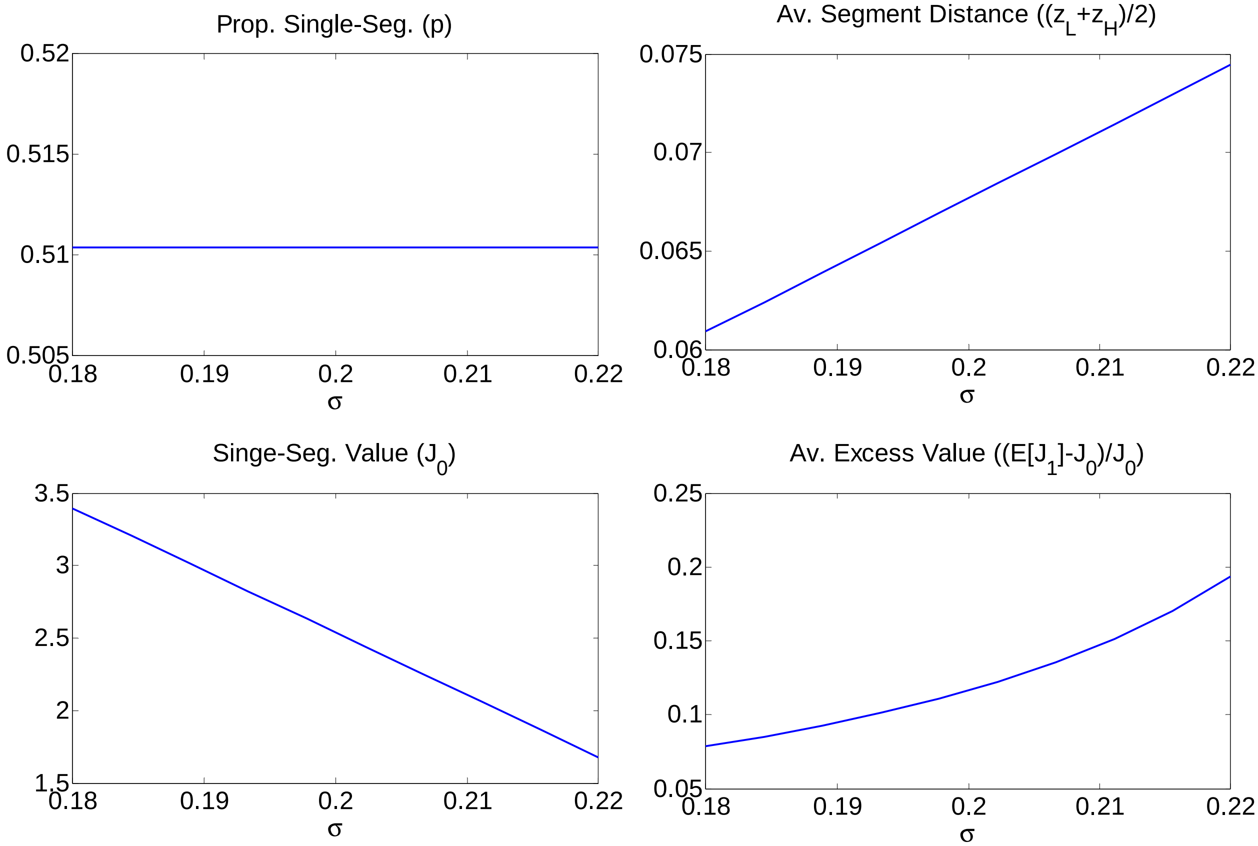

Figure 8: Varying σ: Model Outputs (2/2). The figure shows four key equilibrium outcomes of the model, for σ in an interval of $\pm 10\%$ around the main calibration choice, and no HQ costs ($\beta_0 = 0$); all other parameters are kept at their calibration levels (table 2). The top-left panel shows the proportion of single-segment assets in the economy; the top-right panel shows the average diversified-firm segment distance; the bottom-left panel plots the value of single-segment firms; and the bottom-right panel plots conglomerate excess value.

proposition 4, our model predicts that there exists a strictly positive threshold for σ such that conglomerates fully disappear from the economy. The top-right panel shows how a decrease in σ leads to a decrease in segment distance for the average conglomerate, also in line with data. The result is not that surprising, since a lower σ implies that firms are conducting their business (including M&A deals) within a relatively "tighter" neigborhood. The bottom-left panel shows that the value of single-segment firms increases as σ is reduced, which follows directly from the fact that σ gages the level of project-technology-fit uncertainty. We note in particular that as σ approaches 0, single-segment profits attain their maximum level of 1. Finally, the bottom-right panel of figure 7 shows that excess value increases for higher levels of specialization. This result is more subtle, and actually relies on the existence of an average discount in the economy. Recall that excess value EV corresponds to

$$\frac{\mathrm{E}[J_1] - J_0}{J_0}.$$

Therefore, the derivative of excess value with respect to σ will be negative—as in the bottom-right panel of figure 7—as long as the following holds:

$$\frac{\partial EV}{\partial \sigma} < 0 \Leftrightarrow \frac{\frac{\partial \mathrm{E}[J_1]}{\partial \sigma}}{\frac{\partial J_0}{\partial \sigma}} > \frac{\mathrm{E}[J_1]}{J_0}, \tag{21}$$

where we made use of the fact that $\partial J_0/\partial \sigma < 0$. It turns out that the sensitivity of single-segment firm value to σ is higher than that of diversified-firm value, i.e.,

$$\frac{\partial \mathrm{E}[J_1]}{\partial \sigma} < \frac{\partial J_0}{\partial \sigma}. \tag{22}$$

Whereas we present the above inequality here without an analytical proof, it is true for our numerical examples and it is also intuitive: diversified firms are more prepared to deal with project-technology-fit uncertainty (that is why diversified firms exist), and so variations in σ are less relevant for these corporations relative to their single-segment counterparts. Taken together, equations (21) and (22) imply that excess value will decrease with σ only if

$$\frac{\mathrm{E}[J_1]}{J_0} < 1,$$

which is another way to state that there is an average discount in the economy. In short, our model suggests that there is a connection between the level of the diversification discount and its trend. Pursuing this discussion further, let us inspect figure 8, where we turned off HQ costs, which implies a diversification premium. In this economy, and as argued above, the bottom-right panel shows the opposite relationship between excess value and σ. This economy is also counter-factual with respect to the evolution of the proportion of single-segment assets, which the top-left panel shows is constant. This result is general, as we have shown before in proposition 4. The mechanism behind this result is that without HQ costs profits for both single-segment and diversified firms are linear in the product $\phi\sigma$. Therefore, a decrease in σ does impact profits and valuation, but not the relative preference/advantage of one organizational form *versus* the other.

Taking together model outputs for the cases with and without HQ costs, and in light of data, one is led to the following idea: Extra organizational costs are important in conglomerates (as previous literature suggested), and these costs are not decreasing at a significant rate, at least as compared to gains in specialization.

5 Conclusion

Our paper contributes to the literature by proposing a novel approach to the study of conglomerates, where there is a direct connection between an economy's level of technological specialization and the observed patterns of corporate-diversification activity. Calibrating our model to U.S. data, we are able to match the proportion of assets allocated to conglomerates, the diversification discount, and a positive association between diversified-firm value and an industry-network measure of technological complementarity. The model also shows how an increase in technological specialization can explain aggregate corporate-diversification trends.

Appendix

A.1 Empirical variable definitions

- *Assets*: The total assets of a company (Source: AT variable in COMPUSTAT).
- *Capex*: Funds used for additions to PP&E, excluding amounts arising from acquisitions (Source: CAPEX variable in COMPUSTAT).
- *EBIT (Earnings Before Interest and Taxes)*: Net Sales, minus Cost of Goods Sold minus Selling, General & Administrative Expenses minus Depreciation and Amortization (Source: EBIT variable in COMPUSTAT).
- *Excess Assets*: The log-difference between the assets of a conglomerate and the assets of a similar portfolio of single-segment firms. (Source: COMPUSTAT Segment and Authors Calculations).
- *Excess Capex/Sales*: The difference between the capex/sales of a conglomerate and the capex/sales of a similar portfolio of single-segment firms. We did not take the log difference as in other excess measures because in a few cases Capex/Sales is negative (Source: COMPUSTAT Segment and Authors Calculations).
- *Excess Centrality*: The log-difference between the closeness centrality of a conglomerate and the assets-weighted closeness centrality of a similar portfolio of single-segment firms, using the detailed Input-Output industry classification system (Source: COMPUSTAT, COMPUSTAT SEGMENTS, BEA, and Authors Calculations).
- *Excess EBIT/Sales*: The difference between the EBIT/sales of a conglomerate and the EBIT/sales of a similar portfolio of single-segment firms. We did not take the log difference as in other excess measures because in many cases EBIT/Sales is negative (Source: COMPUSTAT Segment and Authors Calculations).

- *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 single-segment firms, using the detailed 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 detailed Input-Output industry classification system (Source: COMPUSTAT SEGMENTS and BEA).

- *Related Segments*: The number of unique segments of a conglomerate using the detailed Input-Output industry classification system, minus the number of unique segments of a conglomerate using the 3-digit Input-Output industry classification system, following Berger and Ofek (1995) (Source: COMPUSTAT SEGMENTS and BEA).

- *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). We scale the raw variable by its unconditional mean.

- *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).

- *Vertical Relatedness*: Constructed following Fan and Lang (2000). Measures the average input-output flow intensity between each of the conglomerate's non-primary segments and the conglomerate's primary segment; averaged across all non-primary segments. (Source: COMPUSTAT SEGMENTS, BEA, and Authors' Calculations).

A.2 Summary statistics of dataset

Table A.1: Summary Statistics. The table presents summary statistics for each variable. All variables are defined in detail in section A.1.

Panel A: Conglomerates

Variable	Mean	Std. Dev.	Min.	Max.	#Obs.
Tobin's Q	1.682	1.631	0.499	35.16	27,544
Excess Value	-0.284	0.668	-3.062	6.816	27,457
Segment Distance	1.000	0.560	0.046	4.371	27,544
Excess Centrality	0.160	0.109	0.006	0.934	27,544
Vert. Relatedness	18.484	50.136	0	462.8	27,544
N. Segments	2.613	0.937	2	10	27,544
Related Segments	0.345	0.639	0	6	27,544
Assets	4,809	15,533	0.081	340,647	27,544
EBIT/Sales	-0.150	8.925	-1,018	642.3	26,766
Capex/Sales	0.134	2.963	-0.940	433.1	27,206
Excess Assets	-0.105	2.352	-10.861	10.459	27,457
Excess EBIT/Sales	2.829	15.13	-1,018	650.0	26,668
Excess Capex/Sales	-0.707	6.940	-282.5	433.0	27,114

Panel B: Single-Segment Firms

Variable	Mean	Std. Dev.	Min.	Max.	#Obs.
Tobin's Q	2.572	3.271	0.499	35.193	98,564
Assets	1,875	23,403	0.001	3,221,972	119,588
EBIT/Sales	-6.410	165.9	-28,838	5,638	111,441
Capex/Sales	1.180	46.11	-693.2	7,826	117,656

A.3 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 σ 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.2.

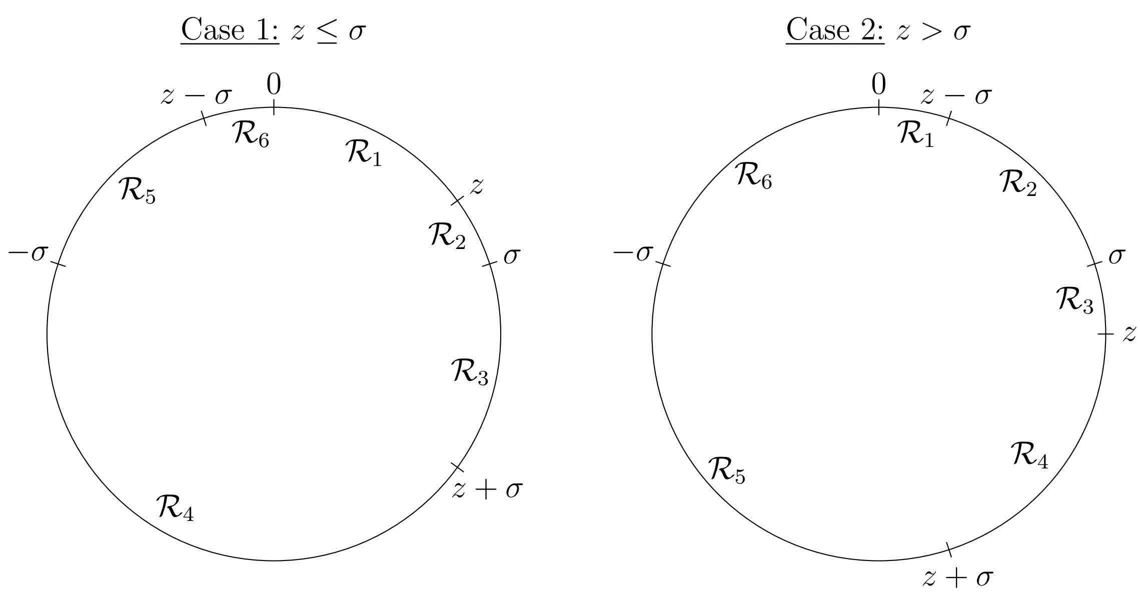

Figure A.1: Splitting the circle into regions. In the left example, $\sigma = 0.2$ and $z = 0.15$. In the right example, $\sigma = 0.2$ and $z = 0.25$.

Location of α_{P_i}	Location of α_{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.2: Optimal allocation policy (swap/no-swap) when two projects occur, as a function of project location; with $z \leq \sigma$.

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

$$\begin{aligned}
&\mathrm{E}[z_{i,P_i^*}] = \\
&= \Pr\{\alpha_{P_i} \in \mathcal{R}_1\}\Big[\Pr\{\alpha_{P_j} \in \mathcal{R}_1\}\mathrm{E}[\min(z_{i,P_i}, z_{i,P_j})|\alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_1] + \\
&+ \Pr\{\alpha_{P_j} \in \mathcal{R}_6\}\mathrm{E}[z_{i,P_j}|\alpha_{P_j} \in \mathcal{R}_6] + \big(1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}\big)\,\mathrm{E}[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_1]\Big] + \\
&+ \Pr\{\alpha_{P_i} \in \mathcal{R}_2\}\Big[\Pr\{\alpha_{P_j} \in \mathcal{R}_1\}\mathrm{E}[z_{i,P_j}|\alpha_{P_j} \in \mathcal{R}_1] + \\
&+ \Pr\{\alpha_{P_j} \in \mathcal{R}_6\}\mathrm{E}[z_{i,P_j}|\alpha_{P_j} \in \mathcal{R}_6] + \big(1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}\big)\,\mathrm{E}[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_2]\Big] + \\
&+ \Pr\{\alpha_{P_i} \in \mathcal{R}_5\}\mathrm{E}[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_5] + \Pr\{\alpha_{P_i} \in \mathcal{R}_6\}\mathrm{E}[z_{i,P_i}|\alpha_{P_i} \in \mathcal{R}_6]. \qquad \text{(A.1)}
\end{aligned}$$

The expression (as a function of parameters) of each of the components in equation (A.1) is presented in table A.3.

We are omitting the explicit integration procedures, since all conditional distributions are

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

Table A.3: Auxiliary table for derivation of equation (A.2).

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 $\mathrm{E}[\min(z_{i,P_i}, z_{j,P_j})|...]$, where we used a standard result on order statistics for random variables drawn from independent uniform distributions.[A.1]

Inserting the expressions from table A.3 into equation (A.1), and after a few steps of algebra, one obtains

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

which implies equation (3) in the proposition.

Case 2: $z > \sigma$

[A.1] 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 $\mathrm{E}[\min(z_{i,P_i}, z_{j,P_j})|...] = z/3$.

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 (5) 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 σ, 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.4 shows the optimal allocation policy for each scenario.

Location of α_{P_i}	Location of α_{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.4: Optimal allocation policy (swap/no-swap) when two projects occur, as a function of project location; with $z > \sigma$.

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

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

The expression of each of the components in equation (A.3) is presented in table A.5. Inserting the expressions from table A.5 into equation (A.3), and after a few steps of algebra,

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

Table A.5: Auxiliary table for derivation of equation (A.4).

one obtains

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

which implies expression (4) in the proposition.■

Proof of proposition 2.

Let us start by conjecturing that the optimal segment distance is smaller than σ. Then we need to obtain the first-order condition with respect to equation (3), 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 (6) 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

(4) 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 (9) is a simple application of 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^* if z_{max} is not binding (where $\pi_1^G(z_L)=\pi_1^G(z_H)$); otherwise, optimal policies take the form $[z_L, z_H=z_{max}]$ (where $\pi_1^G(z_L)<\pi_1^G(z_H)$).
The solution to the problem described in expression (9), given financial markets' equilibrium, needs to verify the following conditions:

$$\begin{cases} rJ_1(z,t,\tau)\,dt=\left[\pi_1^G(z)-\beta_0e^{\beta_1(t-\tau)}\right]\,dt+\mathrm{E}_t[dJ_t], & \text{(A.5)}\\ rJ_0\,dt=\pi_0\,dt+\mathrm{E}_t[dJ_t] & \text{(A.6)}\end{cases}$$

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

$$\begin{aligned} &rJ_1(z,t)\,dt=\pi_1^G(z)\,dt+\lambda_1\,dt\,[J_0-J_1(z,t)]+(1-\lambda_1\,dt)\,\frac{\partial J_1(z,t)}{\partial t}\,dt \Leftrightarrow\\ &J_1(z,t)(r+\lambda_1)-\frac{\partial J_1(z,t)}{\partial t}+\beta_0e^{\beta_1 t}-\left(\pi_1^G(z)+\lambda_1J_0\right)=0 \qquad \text{(A.7)}\end{aligned}$$

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

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

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:

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

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:

$$rJ_0\,dt = \pi_0\;dt + \lambda_0\,dt\,q\,\{\mathrm{E}[J_1(z,t+dt)|z \in [\underline{z},\overline{z}]] - J_0\} + (1 - \lambda_0\,dt\,q)\,0\,dt \Leftrightarrow$$
$$J_0(r + \lambda_0\,q) = \pi_0 + \lambda_0\,q\,\{\mathrm{E}[J_1(z,t+dt)|z \in [\underline{z},\overline{z}]]\} \Leftrightarrow$$
$$J_0 = \frac{1}{r + \lambda_0\,q}\left(\pi_0 + \lambda_0\,q\,\{\mathrm{E}[J_1(z,t+dt)|z \in [\underline{z},\overline{z}]]\}\right) \tag{A.9}$$

where q is the probability of merger acceptance, conditional on a match taking place, defined in the proposition. Combining equation (A.9) and equation (10) yields expression (11).■

Proof of proposition 4.

Let us begin with the first result in the proposition. Since in equilibrium the distribution of firms is stationary, it needs to be the case that the mass of single-segment firms becoming diversified over an infinitesimal dt, $p\lambda_0 q\,dt$, 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. Next we turn to the second result of the proposition, and let us start with the sufficiency argument. If $q = 0$ then no single-segment 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 (10). Replacing π_0 and $\pi_1^G(z^*)$ by their expressions as a function of primitives σ and ϕ (equations (2) and (7)); 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 6 an equilibrium always exists, it must be the case that $q > 0$ holds in equilibrium.

Finally we turn to the third result in the proposition, and let us conjecture that the result is correct, and consider the case where $z_H \leq \sigma$. After some steps of algebra one can then write expected gross profit as

$$\pi_1^G = 1 - \phi\frac{\sigma}{2} + \frac{\phi\sigma}{\theta_H - \theta_L}\left(\frac{\theta_H^4 - \theta_L^4}{96} - \frac{\theta_H^3 - \theta_L^3}{12} + \frac{\theta_H^2 - \theta_L^2}{8}\right), \tag{A.10}$$

where we used expressions (3) and (13). Next, using the fact that in equilibrium $J_0 = \pi_1^G(z_L)/r$, and using equation (11) with $\beta_0 = 0$, we can write

$$\pi_1^G(z_L) = \frac{r+\lambda_1}{r+\lambda_1+\lambda_0 q}\pi_0 + \frac{\lambda_0 q}{r+\lambda_1+\lambda_0 q}\overline{\pi}_1^G.$$

Employing equation (A.10) and the fact that under the conjecture we have $q = (\theta_H - \theta_L)/\eta$, after some algebra the above equality reduces to

$$\frac{\theta_L^3}{24} - \frac{\theta_L^2}{4} + \frac{\theta_L}{4} = \frac{\lambda_0}{\eta(r+\lambda_1) + \lambda_0(\theta_H - \theta_L)}\left(\frac{\theta_H^4 - \theta_L^4}{96} - \frac{\theta_H^3 - \theta_L^3}{12} + \frac{\theta_H^2 - \theta_L^2}{8}\right),$$

which does not depend on either ϕ or σ. Moreover, it does not matter for the proof whether z_{max} is binding or not, since in any case z_H is still just σ multiplied by a constant. The proof for the case where $z_H > \sigma$ follows similar steps, and is therefore omitted.■

Proof of proposition 5.

For all diversified firms operating at segment distance z, the distribution of ages follows a negative exponential distribution with parameter λ_1 (the break-up rate of conglomerates). It follows that

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

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

Proof of proposition 6.

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 single-segment (i.e. $p=1$). Next let us establish that an equilibrium always exists for $\phi\sigma > W$. Taking z_H as given, the optimality condition for z_L 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 such that

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

Using expressions (11) and (10), further manipulation yields

$$\frac{\pi_1^G(z_L)}{r+\lambda_1} - \frac{\beta_0}{r+\lambda_1-\beta_1} = J_0\left(1-\frac{\lambda_1}{r+\lambda_1}\right) \Leftrightarrow \pi_1^G(z_L) - \frac{\beta_0}{1-\frac{\beta_1}{r+\lambda_1}} = J_0 r \Leftrightarrow$$

$$\pi_1^G(z_L) - \frac{\beta_0}{1-\frac{\beta_1}{r+\lambda_1}} = \frac{r+\lambda_1}{r+\lambda_1+\lambda_0 q}\pi_0 + \frac{\lambda_0 q}{r+\lambda_1+\lambda_0 q}\left(\overline{\pi}_1^G - \beta_0\frac{1}{1-\frac{\beta_1}{r+\lambda_1}}\right) \quad \text{(A.11)}$$

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.11) as

$$\begin{aligned}
\pi_1^G(z_L) - \frac{\beta_0}{1 - \frac{\beta_1}{r+\lambda_1}} &= \frac{r+\lambda_1}{r+\lambda_1+\lambda_0 q}\left(\pi_1^G(z_0) - \beta_0\frac{1}{1-\frac{\beta_1}{r+\lambda_1}}\right) + \\
&+ \frac{\lambda_0 q}{r+\lambda_1+\lambda_0 q}\left(\overline{\pi}_1^G - \beta_0\frac{1}{1-\frac{\beta_1}{r+\lambda_1}}\right) \Leftrightarrow \\
\pi_1^G(z_L) &= \frac{r+\lambda_1}{r+\lambda_1+\lambda_0 q}\pi_1^G(z_0) + \frac{\lambda_0 q}{r+\lambda_1+\lambda_0 q}\overline{\pi}_1^G \qquad \text{(A.12)}
\end{aligned}$$

Noting that $\pi_1^G(z_L) \leq \overline{\pi}_1^G$ (since $\pi_1^G(z_H) \geq \pi_1^G(z_L)$), then continuity implies the existence of $z_L \in [z_0, z_H]$ that satisfies equation (A.12). Next we turn to z_H, which given z_L is pinned down uniquely either by the restriction $z_H \leq z_{max}$, or, if this condition is not binding, by the equality

$$\pi_1^G(z_H) = \pi_1^G(z_L),$$

where we note that in such case it needs to be true that $z_H > z^*$. Finally, uniqueness follows from continuity and the fact that the equilibrium is unique at $\phi\sigma \leq W$ (see for example Garcia and Zangwill, 1982 for more technical details).■

References

Ahern, Kenneth R., and Jarrad Harford, 2012, The importance of industry links in merger waves, *Journal of Finance* (forthcoming).

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, 2013, 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.

Bena, Jan, and Kai Li, 2013, Corporate innovations and mergers and acquisitions, *Journal of Finance* (forthcoming).

Berger, Philip G., and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–65.

Bernardo, Antonio E., and Bhagwan Chowdhry, 2002, Resources, real options, and corporate strategy, *Journal of Financial Economics* 63, 211–234.

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.

Diamond, Peter A., 1993, Search, sticky prices, and inflation, *Review of Economic Studies* 60, 53–68.

Dixit, Avinash, and Victor Norman, 1980, *Theory of International Trade* (Cambridge University Press).

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: the impact of prices on takeovers, *Journal of Finance* 67, 933–971.

Fan, Joseph, and Larry Lang, 2000, The measurement of relatedness: An application to corporate diversification, *Journal of Business* 73, 629–60.

Garcia, L.B., and W.I. Zangwill, 1982, *Pathways to Solutions, Fixed Points, and Equilibria* (Prentice-Hall).

Gomes, Joao, and Dmitry Livdan, 2004, Optimal diversification: reconciling theory and evidence, *Journal of Finance* 59, 505–535.

Gormley, Todd A., and David A. Matsa, 2013, Common errors: How to (and not to) control for unobserved heterogeneity, *Review of Financial Studies (Forthcoming)* .

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.

Maksimovic, Vojislav, and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently?, *Journal of Finance* 57, 721–767.

Maskin, Eric, and Jean Tirole, 2001, Markov perfect equilibrium I: Observable actions, *Journal of Economic Theory* 100, 191–219.

Matsusaka, John G., 2001, Corporate diversification, value maximization and organizational capabilities, *Journal of Business* 74, 409–431.

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.

Rajan, Raghuram G., Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance* 55, 35–79.

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.

Ricardo, David, 1817, *The Principle of Political Economy and Taxation* (Gearney Press (1973)).

Sanzhar, Sergey V., 2006, Discounted but not diversified: Organizational structure and conglomerate discount, Working Paper, available at SSRN.

Scharfstein, David S., Robert Gertner, and Eric Powers, 2002, Learning about internal capital markets from corporate spinoffs, *Journal of Finance* 57, 2479–2506.

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.

Smith, Adam, 1776, *An Inquiry into the Nature and Causes of the Wealth of Nations* (Reprint, University of Chicago Press (1976)).

Stein, Jeremy C., 1997, Internal capital markets and the competition for corporate resources, *Journal of Finance* 52, 111–133.

Yang, Xiaokai, and Siang Ng, 1998, *Specialization and division of labor: a survey*, chapter *in* Increasing Returns and Economic Analisys (McMillan).