Technological Specialization and the Decline of Diversified Firms*

Fernando Anjos† Cesare Fracassi‡

Abstract

We document a strong decline in corporate-diversification activity since the late 1970's, and we develop a dynamic model that explains this pattern, both qualitatively and quantitatively. The key feature of the model is that synergies endogenously decline with technological specialization, leading to fewer diversified firms in equilibrium. The model further predicts that segments inside a conglomerate should become more related over time, which is consistent with the data. Finally, the calibrated model also matches other empirical magnitudes well: output growth rate, market-to-book ratios, diversification discount, frequency and returns of diversifying mergers, and frequency of refocusing activity.

May 22, 2015

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

Keywords: corporate diversification, specialization, mergers, matching.

---

*The authors thank comments from and discussions with Kenneth Ahern, Andres Almazan, Aydoğan Alti, Cláudia Custódio, Vojislav Maksimovic (Tepper/LAEF discussant), Matt Rhodes-Kropf (AFA discussant), Alessio Saretto, and Laura Starks. The authors also thank comments from seminar participants at the University of Texas at Austin and the NOVA School of Business and Economics, and participants at the following conferences: 2012 European meetings of the Econometric Society, 2013 North American Summer meetings of the Econometric Society, 2014 meetings of the American Finance Association, and 2014 Tepper/LAEF Macro-Finance conference.

†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

Much finance research on corporate diversification has focused on the cross section of firms, in an attempt to understand whether and how conglomerates create value. However, little attention has been devoted to studying corporate-diversification time trends, the topic of our paper. We start by documenting a steady decline in corporate diversification since the late 1970's. We then develop and calibrate a dynamic model that accounts for this trend, as well as several other empirical magnitudes associated with corporate diversification.

First, we present novel evidence about the decline of corporate diversification activity in the United States: over the last 35 years, we document (i) a decrease in the number of conglomerates (from 55% of the total number of public firms to around 25%), (ii) a decrease in the number of segments (or industry-level divisions) for the average conglomerate (from 3.2 to 2.6), (iii) an increase in the importance of the conglomerate's main segment (from 64% of total assets to 72%), and (iv) a decrease in diversifying merger activity (around a 14 percentage points decline in the fraction of total mergers).

Our theory for why conglomerates decline builds on the seminal economic concept of *technological specialization*, or division of labor. An ever-increasing technological specialization was the mechanism proposed by early authors as the key driver of economic growth (Smith, 1776; Ricardo, 1817), and these ideas were later formalized by various economists (e.g., Rosen, 1978; Yang and Borland, 1991; Becker and Murphy, 1994).[1] In our model, technological specialization refers to the quality of the match between a worker's skills and the task she is assigned to, and we assume that specialization increases at an exogenous rate. For example, such a trend can be thought of as the product of better information and communication technologies, as argued for instance in Varian (2010): "*[...] communications technology allows tasks to be modularized and touted to the workers best able to perform those tasks.*"

In our model, technological specialization interacts importantly with the mechanism that generates synergies from corporate diversification. Specifically, a diversified firm is an organization that employs workers with a more diverse set of skills, and where workers can exchange tasks amongst themselves if this improves the efficiency of the task-skill match. In times of greater technological specialization, it is more likely that the original task-skill match is already relatively efficient, thus there is less scope for gains from within-firm resource reallocation. Thus the model delivers the implication that in equilibrium there are

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

fewer conglomerates over time. Such a negative relationship between technological specialization and corporate diversification is not *a priori* obvious. For example, if coinsurance benefits were the main rationale for the existence of conglomerates, and if more specialized business units had less correlated payoffs, then specialization would favor the conglomerate form.

We acknowledge that other explanations for the decline of corporate diversification are possible, and our theory does not exclude that additional mechanisms might be at play. For example, corporate governance has been improving over the last decades, which could gradually reduce the presence of conglomerates that are motivated by empire-building motives, as argued by Denis, Denis, and Sarin (1997). However, we note that ours is a parsimonious model that (i) accounts for several observed empirical patterns relatively well, as detailed below; and (ii) builds on a seminal concept from economics, namely technological specialization.

We model an economy that is populated by a continuum of business units, which can be thought of as collections of workers with relatively homogeneous skills. Time is continuous, and single-segment firms can engage in diversifying mergers.[2] Following Rhodes-Kropf and Robinson (2008), mergers are modeled in the spirit of search-and-matching labor economics (Diamond, 1993; Mortensen and Pissarides, 1994): single-segment firms meet up at random according to an exogenous Poisson process, and then decide whether to become a conglomerate. Diversification synergies are positive when a conglomerate is initially formed, but with some probability the conglomerate becomes inefficient, incurring additional overhead costs. Once a conglomerate becomes inefficient, it refocuses with some probability, also according to an exogenous Poisson process.

We model production technology, specialization, and diversification synergies using a spatial representation. Specifically, each business unit faces a project opportunity (alternatively, a collection of tasks required for production), and both the business unit and the project are characterized by a location on a technology circle. The location of the business unit refers to the core technological skill of its workers, and output decreases in the

[2]For simplicity, corporate diversification and refocusing in our model are entirely driven by mergers and spin-offs. The assumption of focusing on corporate-restructuring mechanisms is consistent with previous literature: almost two thirds of the firms that increase the number of segments implement this strategy via acquisition (Graham, Lemmon, and Wolf, 2002); and many diversifying mergers are later divested (Ravenscraft and Scherer, 1987; Kaplan and Weisbach, 1992; Campa and Kedia, 2002).

distance between the business unit and the project. A key feature of the model is that project location is uncertain, and thus business units face the risk of drawing a project for which they are ill-equipped, which motivates corporate diversification. As explained before, diversification generates synergies because business units within the same firm are allowed to trade projects—or, equivalently, reallocate resources—whenever this is efficient. Such resource reallocation is not available across firms, which can be motivated by the existence of informational frictions and/or coordination problems.[3]

In our spatial approach, we interpret the range of project locations faced by each business unit as the degree of technological specialization. In periods of low specialization this range is wide, which implies corporate diversification can significantly add value via frequent and effective ex-post reallocation. As specialization increases, business units generally face projects for which they have a comparative advantage, with two implications: average output increases and diversification synergies decrease, which leads to fewer conglomerates in equilibrium. Both these predictions are consistent with data.

In our model all conglomerates have two segments, located at a certain distance in the technology circle. The model implies that there is an interior optimal technological distance between segments, driven by the following trade-off. On the one hand, complementarity is relatively low if two business units are technologically very similar, since resource reallocation only generates limited gains. We thus would expect that diversifying synergies initially increase in technological distance between segments. On the other hand, if segment distance is too high, there are very few opportunities for reallocation. A key implication of our model is that optimal segment distance decreases with technological specialization, since a more-focused business unit requires a relatively closer counterpart for efficient within-firm reallocation to take place. This prediction is consistent with the observed trend for the average level of relatedness across segments: our main empirical relatedness measure decreases by about 15% over the period 1990-2013.

Using data on corporate-diversification activity in the U.S., we then perform a calibration of our dynamic model. The calibration employs a growth rate for technological specialization that generates reasonable output growth, and we use six other empirical moments to identify

[3]This rationale is consistent with interpreting the boundaries of the firm as information boundaries, as suggested for example in Chou (2007). Informational frictions also play a prominent role in certain theories of the firm, in particular transaction-cost economics (Coase, 1937; Williamson, 1975).

the model's remaining parameters: (i) the fraction of assets allocated to single-segment firms in the economy, (ii) the level of the market-to-book ratio, (iii) the level of the diversification discount, (iv) the likelihood that a firm engages in M&A, (v) average diversifying-merger announcement returns, and (vi) the average rate at which conglomerates refocus. The model is able to match these moments fairly well, but, more importantly, it also matches several key magnitudes that had no direct bearing in the calibration: (i) the rate at which conglomerates are declining, (ii) the rate at which single-segment market-to-book ratio is increasing, and (iii) a relatively flat diversification discount.

One of the interesting features of the model is that we can make predictions about the future evolution of corporate diversification. According to our calibration, diversifying mergers will cease by the early 2050's and conglomerates will represent only about 1% of the total assets in the economy by the end of this century, compared to about 54% at the end of 2013.

One of the critical features of the model is segment distance, defined as the technological distance across divisions. As mentioned before, the model predicts that the average segment distance decreases over time. In order to test this implication, we introduce a novel empirical measure of cross-division relatedness, which, in the spirit of the model, also employs a spatial approach. Specifically, we follow Acemoglu et al. (2012), Ahern and Harford (2014), and Anjos and Fracassi (2015) and construct an inter-industry network using input-output flows. With such network we can compute the average distance across conglomerate segments, by taking into account all direct and indirect inter-industry relationships in the economy. Using this measure, segment distance decreases by about 15% over the period 1990-2013, a trend which the calibrated model matches almost perfectly, even though it had no direct bearing in parameter choice.

We further investigate whether the model can account for cross-sectional relatedness patterns. First we find that conglomerates cluster at intermediate segment distances, which is consistent with the model's prediction about the existence of an interior optimal segment distance. Second, we find a positive association between segment distance and conglomerate value. This association does not match the non-monotonic implication from the model, possibly because of adverse-selection concerns that are more serious for distant mergers. In the appendix, we provide an extension to our main model that accounts for the observed relationship between segment distance and conglomerate value.

Our paper mostly relates to finance literature on corporate diversification. Starting with two seminal empirical papers (Lang and Stulz, 1994; Berger and Ofek, 1995), financial economists have asked whether conglomerates trade at a discount, when compared to benchmark portfolios of single-segment firms. Both these papers found significant diversification discounts,[4] which would be consistent with explanations emphasizing the "dark side" of conglomerates (Scharfstein and Stein, 2000; Scharfstein, Gertner, and Powers, 2002; Rajan, Servaes, and Zingales, 2000). Our model partly draws on this literature in that there exists a sizable cost associated with organizational complexity (not incurred by single-segment firms). However, ours is a trade-off model of diversification, where we simultaneously consider costs and benefits to this activity. Moreover, we introduce a new framework for the "bright side" of corporate diversification, one that emphasizes the role of resource reallocation and technological specialization. This approach expands on previous literature on the advantages of internal capital markets, where conglomerate headquarters potentially reallocate capital from low-productivity to high-productivity divisions (Stein, 1997; Hubbard and Palia, 1999; Scharfstein and Stein, 2000; Maksimovic and Phillips, 2002).[5]

Our dynamic approach to modeling corporate diversification follows in the footsteps of several other papers (Matsusaka, 2001; Bernardo and Chowdhry, 2002; Gomes and Livdan, 2004). Our paper is different in that we emphasize the role of technological specialization in determining synergies; and, furthermore, in that we focus on explaining corporate-diversification trends.

Finally, our paper has methodological similarities with other dynamic approaches to M&A (Yang, 2008; Hackbarth and Morellec, 2008; Morellec and Zdhanov, 2008; David, 2014; Dimopoulos and Sacchetto, 2014), however, none of these papers focuses on the topic of corporate diversification.

1 The evolution of corporate diversification

We begin by documenting a set of comprehensive corporate-diversification trends in the United States over the last 35 years, which are depicted in figures 1 and 2. Figure 1 shows

[4] The discount discovered in Lang and Stulz (1994) and Berger and Ofek (1995) has been challenged by much subsequent empirical research. See, for example, Custódio (2014).

[5] Also see a recent paper on the benefits of internal labor markets (Tate and Yang, 2015) and a recent paper on capital and labor reallocation within firms (Giroud and Mueller, 2015).

four different measures of corporate-diversification activity: average number of segments in a conglomerate (row 1), fraction of assets allocated to a conglomerate's main segment (row 2), fraction of assets in the economy allocated to single-segment firms (row 3), and finally fraction of firms in the economy that are single-segment (row 4). For each of these four measures we employ two alternative industry classifications. The first, shown in the left-side panels, is the Input-Output (I-O) industry classification from the 1997 detailed I-O tables. These I-O tables contain cross-industry flows of goods and services for 470 industries and are based on an aggregation of codes from the North American Industry Classification System (NAICS). NAICS codes are available since 1990, hence our NAICS/I-O time series start in 1990. The second industry classification we use is based on the more-standard 4-digit SIC codes, which go back further and allow us to construct time series starting in 1977, the first year we have firm data available from COMPUSTAT Segment.[6] The advantage of using SIC codes is that we have longer time series. The advantages of the NAICS/I-O classification are twofold: (i) it was created more recently, and thus it is ostensibly an industry classification scheme that better describes the actual economy; and (ii) it allows us to construct an I-O-based cross-segment *relatedness* measure that is required for a later analysis (section 4.1). Also, by using two industry classification systems we are showing that key time trends are not driven by the specific choice of industry classification.

The two panels in the first row show that diversified firms have been gradually reducing the number of different industries where they operate, from about 3.2 in 1977 to about 2.6 in 2013 (a 19% decline), using the SIC classification. The second row shows that conglomerates have been allocating a lower proportion of their assets to secondary segments, defined as all segments but the main one: while secondary segments accounted for approximately 37% of all assets in 1977, such lines of business account for only about 28% in 2013, again according to the SIC classification. The panels in the first and two rows thus illustrate that the average conglomerate is becoming more similar to a single-segment firm.

The third- and fourth-row panels turn to a comparison between conglomerates and single-segment firms. A discontinuity is observed in the transition from 1997 to 1998, and this is a consequence of the change in segment-reporting requirements introduced at the end 1997.[7] Focusing on periods before and after the discontinuity, the third- and fourth-row panels show

[6] The year 1976 has only few observations.

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

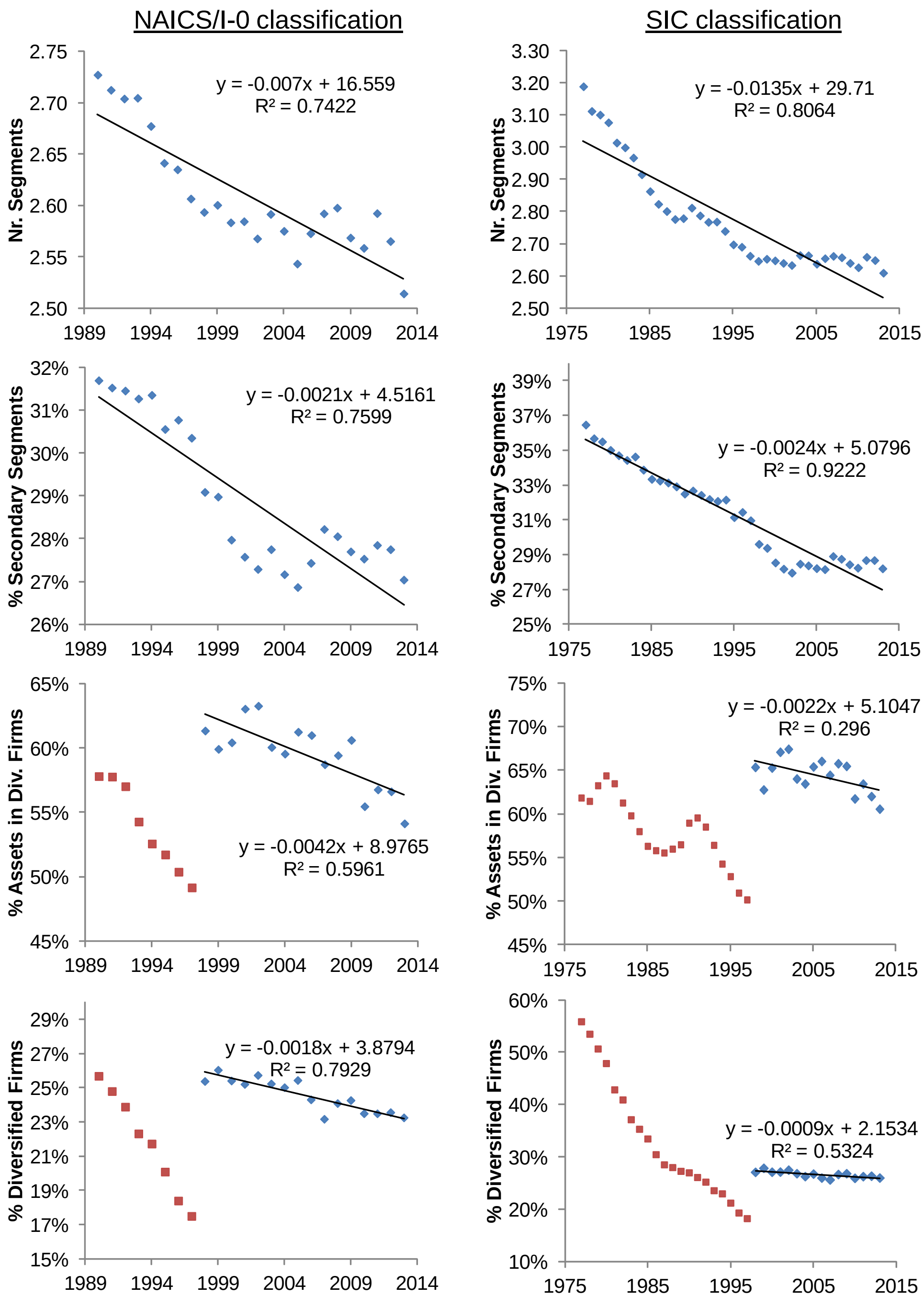

Figure 1: Evolution of Corporate Diversification. The figure shows four measures of corporate diversification in the United States: average number of segments in a conglomerate (row 1), fraction of assets allocated to conglomerate's secondary segments (row 2), fraction of assets in the economy allocated to diversified firms (row 3), and fraction of diversified firms in the economy (row 4). The left panels use the 1997 Input-Output (I-O) classification at the detailed level, which aggregates NAICS codes. The right panels use the 4-digit SIC industry classification.

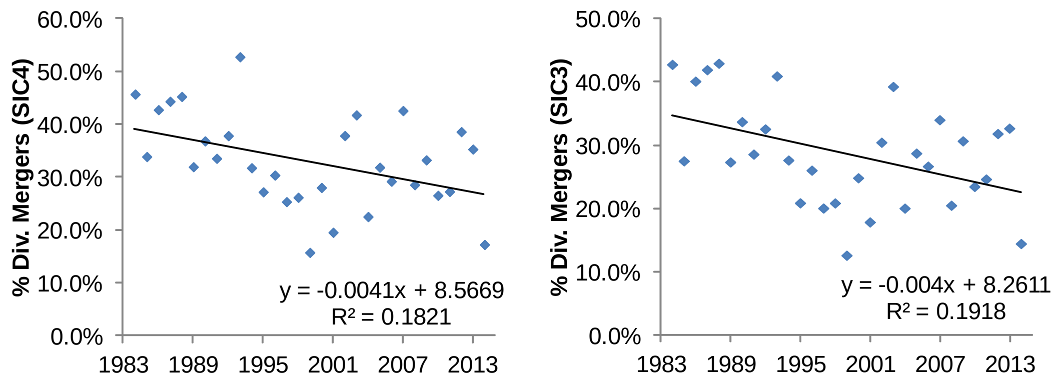

Figure 2: Evolution of Diversifying Mergers. The figure shows the fraction of merger deals (in dollar amount) that are diversifying. In the left (right) panel a diversifying merger is defined as a deal between two firms that have no overlapping SIC 4 (SIC 3) codes.

that there is a decline in the presence of conglomerates in the economy. For example, using the NAICS-IO classification for the period after the discontinuity (third-row left panel), an average of about one percent of the economy's assets shifts from conglomerates to single-segment firms every 2 to 3 years. A similar pattern is seen in the fraction of firms that are diversified (fourth-row panels).

Finally, data on merger activity also supports the view that corporate diversification has been declining: figure 2 shows a steady decline in diversifying mergers, as a fraction of total merger activity (dollar amount), in the order of 0.4 percentage points per year over the last 30 years (i.e., 12 percentage points from 1984 to 2014). To construct the plots, we use domestic US mergers data from Thomson Reuters SDC, for the period 1984-2014, and include public firms, private firms, and subsidiaries. We classify a merger as diversifying if there is no overlap between the SIC codes of the merging entities. The left panel employs SIC codes at the 4-digit level, the right panel at the 3-digit level.

Overall, figures 1 and 2 suggest a long-term decline in corporate diversification. This trend is consistent with findings in the corporate diversification literature that focused on specific time periods. For example, Denis, Denis, and Sarin (1997) show that the average number of segments declined from 2.4 in 1985 to 2.1 in 1989. Comment and Jarrell (1995) find that the proportion of single-segments firms increased from 36% in 1978 to 64% in 1989. The findings of these papers notwithstanding, our results suggests that the decline in corporate diversification is a long run phenomenon and not driven by specific merger waves.

2 Model

In the previous section we presented strong evidence that corporate-diversification activity has been steadily declining over the last 35 years. We now turn to developing our theoretical framework, which will offer an explanation for the observed trend. We start by constructing a static equilibrium model for flow payoffs (section 2.1), which we then embed in a dynamic search-and-matching framework (section 2.2).

2.1 Flow payoffs

The economy comprises a continuum of business units (BUs), which can be thought of as collections of workers with relatively homogeneous technological skills. Each BU i is characterized by a location α_i on a circle with measure 1, represented in figure 3.[8] The different locations on the circle represent different technologies, which enable BUs to pursue profitable project opportunities. Our notion of technology is broad, and includes not only technical capabilities, but also a firm's managerial/organizational know-how.

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 in section 2.2. The next two subsections further characterize the flow payoffs of single-segment and diversified firms.

2.1.1 Single-segment firms

Each BU in the economy undertakes one project, and this project is also characterized by a location in the technology circle, denoted by α_{P_i}. Project 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 find. The support of the distribution for project location corresponds to the dashed arc in figure 3. The higher σ is, the higher the risk that business units are presented with projects for which they are ill-equipped.

[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.

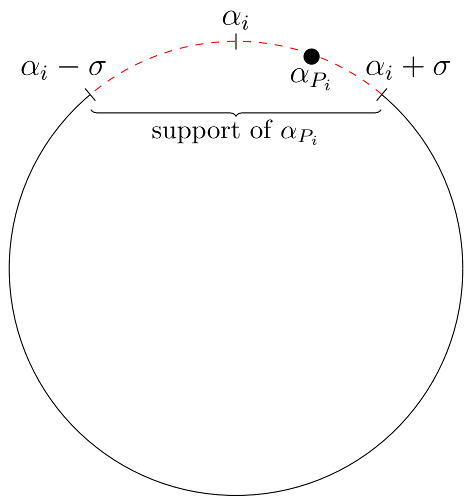

Figure 3: Technologies and Projects: Spatial Representation. The figure depicts a circle where both projects and business units are located. The location of the business unit (α_i) represents its technology and the location of projects (α_{P_i}) represents the ideal technology to undertake that particular project. The figure also shows that business units draw projects from locations close to their technology, in the interval $[\alpha_i, -\sigma, \alpha_i + \sigma]$, where σ is the exogenous level of technological specialization.

We interpret the inverse of σ as the degree of *technological specialization*, which thus refers to the extent to which business units are able to find good projects for their technology. This concept of technological specialization represents the set of institutions and production techniques that enable agents to focus on the specialized set of activities at which they excel, which would lead to higher productivity.

For tractability we assume $\sigma < 1/4$, which simplifies the analysis.[9]

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

[9]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.

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” ($\sigma = 0$). In the dynamic version of the model we assume that σ gradually decreases over time, which thus translates into positive economic growth (dynamics are detailed in section 3.2).

Finally, equation (2) shows that ϕ and σ are not separately identified: as long as the product $\phi\sigma$ is constant, payoffs are the same.[10] This point is important for our calibration, where, given the argument just outlined, we set the initial σ at an arbitrary level.

2.1.2 Diversified firms

To keep the framework tractable, the only form of corporate diversification we consider is a conglomerate with two segments (i.e., two business units). We define *segment distance* as the length of the shortest arc between the two business units in the technology circle, and we denote it as z. As will become apparent shortly, segment distance plays an important role in the economic performance of diversified firms.

If BU i is part of the same firm as BU j, capacity is still assumed to be one project per unit, and thus the profit function is similar to that of a single-segment firm. The key difference is that in conglomerates projects can be traded (swapped) across segments; and this ex-post choice is assumed to be made optimally by the headquarters of the multi-segment firm so as to minimize the total costs of project-technology misfit (represented in figure 4). 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”), the nature of which is imperfectly known ex ante.

An implicit assumption of our model is that projects cannot be traded across firms. This

[10] A caveat is in order. Identification could in principle be obtained under particular assumptions about the matching function that brings single-segment firms together for a potential merger deal and/or the dynamics of σ. However, since the optimal merger distance is in general an increasing function of σ (see proposition 2), such identification would in general be weak and depend on the very specific non-linearities induced by our modeling assumptions.

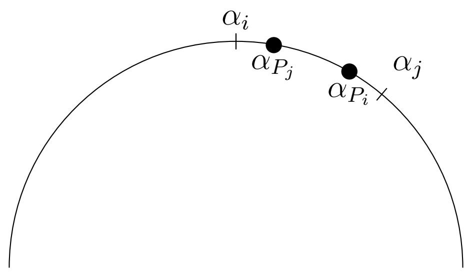

Figure 4: Conglomerates and Reallocation: Spatial Representation. The figure depicts the location of conglomerate segments on the technology circle; and shows an instance where projects are optimally swapped across segments, i.e., division i is assigned to project j and vice-versa.

could be due, for example, to adverse selection; and would be consistent with interpreting the boundaries of the firm as information boundaries (as suggested, e.g., in Chou, 2007).

The economy comprises two types of diversified firms: good conglomerates, which reap the synergistic benefits from diversification at no additional cost; and bad conglomerates, which impose an extra cost on the firm. If bad conglomerates are pervasive enough, the model will imply a diversification discount, as observed in data. For now we take the proportions of good and bad conglomerates as given; these are endogenized later (section 2.2). We first describe the workings of good conglomerates.

Good conglomerates

Below we present the expected profit function for a good conglomerate, taking segment distance in the technology circle as given; these expressions are obtained by computing the likelihood of project transfer and the conditional average gain per transfer (see proof of proposition 1 in the appendix for details).

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

$$\pi_1(z) = \begin{cases} 1 - \phi\dfrac{\sigma}{2} + \phi\left(\dfrac{z^3}{24\sigma^2} - \dfrac{z^2}{4\sigma} + \dfrac{z}{4}\right) & z \leq \sigma \quad \text{(3a)} \\ 1 - \phi\dfrac{\sigma}{2} + \phi\left(-\dfrac{z^3}{24\sigma^2} + \dfrac{z^2}{4\sigma} - \dfrac{z}{2} + \dfrac{\sigma}{3}\right) & \sigma < z \leq 2\sigma \quad \text{(3b)} \\ 1 - \phi\dfrac{\sigma}{2} & z > 2\sigma \quad \text{(3c)} \end{cases}$$

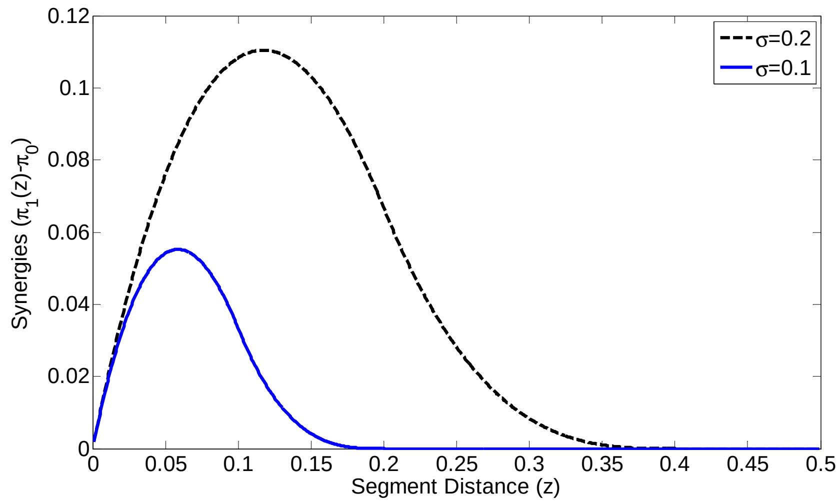

Figure 5: Segment Distance and Synergies. The figure plots diversification synergies as a function of segment distance z. Synergies are the difference between the average divisional payoff of conglomerate, $\pi_1(z)$, and the average payoff of a single-segment firm, π_0. ϕ is set at 8.

Figure 5 depicts the relationship between segment distance z and average synergies, that is, the difference between the average (divisional) payoff of a good conglomerate and the average payoff of a single-segment firm. The figure shows that synergies, holding segment distance z constant, are reduced when σ is lower, i.e., when technological specialization is higher. This occurs because with lower σ there are fewer opportunities for efficient inter-division project trade and, furthermore, each inter-division project trade has a lower average gain. This reduction in synergies can qualitatively explain why over time we observe fewer diversified firms, as documented in section 1 (figure 1). In section 3 we further ask whether the dynamic version of the model can quantitatively account for the observed decline in corporate diversification.

Figure 5 also illustrates how, for constant σ, the relationship between segment distance and average synergies is non-monotonic: if distance is too low, the likelihood of a project transfer is greater, however the average gain of the 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 optimal distance trades off the frequency of desirable transfers with the average gain of each transfer. Proposition 2 shows that the optimal (static) segment distance is a simple proportion of project-type uncertainty σ.

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

$$z^* = \sigma\left(2 - \sqrt{2}\right), \tag{4}$$

with associated expected BU profit of

$$\pi_1(z^*) = 1 - \phi\sigma\left(\frac{2}{3} - \sqrt{\frac{1}{18}}\right). \tag{5}$$

According to proposition 2, an increase in technological specialization (lower σ) implies that diversified firms should become more specialized too, that is, one should observe most conglomerates with lower segment distance (this is also visible in figure 5). As we show later, this is consistent with the patterns we observe in data. Furthermore, we show that the dynamic version of the model can quantitatively match the increase in relatedness (or reduction in segment distance), an analysis we pursue in section 4.1.2.

Finally, it is unclear which *cross-sectional* relationship between segment distance and profits is implied by this simple static model. 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 could be negative. This ambiguity may explain the apparent contradiction between some 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). We further analyze the cross-sectional implications of our model in section 4.1.3.

Bad conglomerates

In our dynamic model, a good conglomerate may become bad at some future point in time, after which each division incurs an additional cost of β. This assumption is consistent with papers on the "dark side" of internal capital markets (Scharfstein and Stein, 2000; Scharfstein, Gertner, and Powers, 2002; Rajan, Servaes, and Zingales, 2000). The extra cost associated with bad conglomerates being independent of segment distance is consistent 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. We impose an assumption relating the level of synergies and the additional overhead β of bad conglomerates:

Assumption 1 *The maximal level of synergies is lower than the additional overhead of bad conglomerates. Formally,*

$$\pi_1(z^*(\phi, \sigma); \phi, \sigma) - \pi_0(\phi, \sigma) < \beta. \tag{6}$$

Assumption 1 implies that it is optimal for any inefficient firm to seek refocusing, which simplifies the analysis of the dynamic model later on. This rationale notwithstanding, the assumption is not binding in our calibration.

2.2 Dynamics

2.2.1 Matching technology

In the previous section we developed a static model for the flow payoffs of diversified and single-segment firms. In this section we embed the flow payoff model in a dynamic framework, which we then fit to data. Specifically, we model 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 tractability 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.

There is an exogenous continuously-compounded discount rate denoted by r and all agents are risk-neutral. Firm boundaries change 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. Modeling diversification as driven by merger and spin-off activity is motivated by the fact that almost two thirds of the firms that increase the number of segments implement this strategy via acquisition (Graham, Lemmon, and Wolf, 2002); and that many diversifying mergers are later divested (Ravenscraft and Scherer, 1987; Kaplan and Weisbach, 1992; Campa and Kedia, 2002).

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 existing single-segment firms is 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. After a conglomerate is formed, it becomes bad according to a Poisson process with intensity λ_1. Under assumption 1 it is efficient to break a bad conglomerate apart. However, we assume there are frictions–such as managerial entrenchment or search costs–to breaking up immediately, and hence refocusing occurs according to a Poisson process, with intensity λ_2.

Finally, we specify that, conditional on a merger opportunity arising, the distance between the two single-segment firms be drawn from a uniform distribution with support $[0, 1/2]$. This assumption is consistent with matched BUs being selected uniformly at random in the technology circle.

2.2.2 Solving the dynamic model: steady-state case

This section solves the model for the particular case where technological specialization is time-invariant, and where we focus on the steady-state equilibrium. Although ultimately we will be calibrating a version of the model where specialization increases over time (i.e., σ decreases over time), the solution to the general case is not analytical. The steady-state case thus provides a useful benchmark to understand the basic mechanics of the model.

We first state the individual optimization problem. Since business units share merger surplus equally, the optimization problem from the perspective of business unit i is as follows:

$$
\begin{aligned}
J_t \;=\; & \sup_{\{\tau\}} \Bigg\{ \mathrm{E}_t \Bigg[\int_{u \in [t,+\infty] \cap \{[\tau,\tau_2]\}} e^{-r(u-t)} \left[\pi_1 \left(z_{\sup\{\tau<u\}} \right) - \beta \mathbb{1}_{\sup\{\tau<u\}<\sup\{\tau_1<u\}} \right] du + \\
& + \int_{u \in [t,+\infty] \setminus \{[\tau,\tau_2]\}} e^{-r(u-t)} \pi_0 \, du \Bigg] \Bigg\}, \qquad (7)
\end{aligned}
$$

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_1 stands for the time at which a good conglomerate formed at τ becomes bad, τ_2 returns the time at which a conglomerate formed at τ splits, and $z_{\sup\{\tau<t\}}$ is the time-t distance of the two divisions inside the diversified firm. The first integral in (7) refers to the present value of cash flows when the BU is operating in a conglomerate, and the second integral refers to the present value of cash flows when the BU is a single-segment firm. When the BU is in a conglomerate, its cash flows are a function of segment distance ($z_{\sup\{\tau<u\}}$) and whether or not the conglomerate is bad ($\beta \mathbb{1}_{\sup\{\tau<u\}<\sup\{\tau_1<u\}}$).

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 fraction of bad conglomerates $w \in [0,1]$, 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 the merger acceptance policy solves optimization problem (7).*

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, denoted by 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 bad conglomerate, J_2, is a simple function of the segment distance at which the merger took place (z):*

$$J_2(z) = \frac{\pi_1(z) - \beta + \lambda_2 J_0}{r + \lambda_2} \tag{8}$$

The value of a business unit inside a good conglomerate, J_1, is given by

$$J_1(z) = \frac{\pi_1(z)(r + \lambda_1 + \lambda_2) - \lambda_1 \beta + \lambda_1 \lambda_2 J_0}{(r + \lambda_1)(r + \lambda_2)}. \tag{9}$$

The value of single-segment firms J_0 is characterized as

$$J_0 = \frac{\pi_0(r + \lambda_1)(r + \lambda_2) + \lambda_0 q(r + \lambda_1 + \lambda_2)\overline{\pi}_1 - \lambda_0 q \lambda_1 \beta}{(r + \lambda_0 q)(r + \lambda_1)(r + \lambda_2) - \lambda_0 q \lambda_1 \lambda_2}, \tag{10}$$

with q the (endogenous) probability of merger acceptance and $\overline{\pi}_1$ the (endogenous) average diversified-BU profit rate of good conglomerates:

$$q := \frac{z_H - z_L}{0.5} \tag{11}$$

$$\overline{\pi}_1 := \int_{z_L}^{z_H} \frac{1}{z_H - z_L} \pi_1(z)\, dz \tag{12}$$

Equation (10) describes the equilibrium value of single-segment firms, which embeds the value of the option to diversify. It is also clear in equations (8)-(10) that the costs associated with bad conglomerates (β) negatively affect equilibrium firm value (including single-segments). 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*

$$p = \frac{1}{1 + \lambda_0 q \left(1/\lambda_1 + 1/\lambda_2\right)}. \tag{13}$$

2. *The fraction of bad conglomerates is*

$$w = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{14}$$

3. *There exists a threshold C, defined as*

$$C := \frac{6\lambda_1 \beta}{(\sqrt{2} - 1)(r + \lambda_1 + \lambda_2)}, \tag{15}$$

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

The first result in proposition 4 shows that, holding the merger acceptance probability constant, the steady-state proportion of single-segment firms increases in both λ_1 and λ_2; and decreases in λ_0. This is intuitive, since higher λ_1 (likelihood of becoming bad conglomerate) or λ_2 (likelihood of bad conglomerate refocusing) speed up the average rate at which a conglomerate ultimately refocuses, and λ_0 determines the frequency of diversifying-merger opportunities.

The second result shows that the fraction of bad conglomerates in equilibrium is entirely driven by the entry-rate/exit-rate ratio of such firms. This implies that if extra overhead costs β incurred by bad conglomerates are large enough and the intensity of refocusing λ_2 is small enough (relative to λ_1), the economy will exhibit an average diversification discount. The discount is due to the long-run (or unconditional) proportion of bad conglomerates is high (these firms rarely break up). Nevertheless, it may still be optimal for single-segment firms to engage in diversifying mergers ex-ante, as long as λ_1 is low as well. The discount is a poor measure of the relative value of diversified firms because it does not take into account the value that was created by bad conglomerates at a previous time where they were still good.[11]

The third result in proposition 4 shows that mergers only take place if either technological specialization is low (high σ) or the cost of project-technology misfit is high (ϕ), relative to organizational costs (β). As derived in the static-setup section, the advantage of a conglomerate is the ability to optimize BU-project assignment ex-post (representing resource reallocation), an option assumed to be unavailable to single-BU firms. These benefits of diversification are compared to its costs, gauged by the parameter β. These costs are less important if only incurred for a short period of time, that is, when λ_2 is high. Finally, when $\lambda_1 \to 0$, organizational-complexity costs no longer factor into the diversification trade-off, since bad conglomerates almost never materialize.

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

Proposition 5 *The equilibrium specified in definition 1 always exists and is unique.*

2.2.3 Solving the dynamic model: non-stationary case

In the appendix we detail the solution to the non-stationary case (section A.2) where technological specialization increases over time (i.e., σ decreases over time) and outline the numerical implementation. In this section we summarize the main difference between the stationary and non-stationary case.

[11] This argument is along the lines of Anjos (2010).

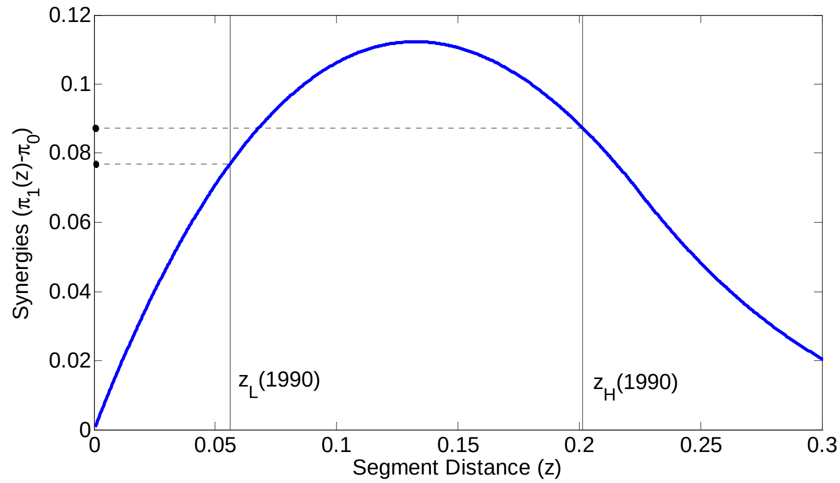

Figure 6: Dynamic Calibration: Policies. The figure shows the optimal merger-acceptance region, delimited by z_L and z_H, and the level of flow synergies as a function of segment distance z (solid curve). Synergies are the difference between the average divisional payoff of a conglomerate, $\pi_1(z)$, and the average payoff of a single-segment firm, π_0. The plot is based on the main calibration, for the year 1990 (section 3.2).

The dynamic model is not simply a comparative statics of the steady-state, where we decrease σ over time. The key difference in the optimization problem is that firms *anticipate* that σ changes over time. As a result they adjust their policies accordingly, and this actually induces a relative preference for low segment distances. This is illustrated in figure 6, which takes a snapshot of our dynamic calibration for the year 1990 (the construction of this calibration is detailed later on).

The vertical lines in figure 6 are the endpoints of the optimal merger-acceptance region $[z_L, z_H]$ at 1990, and the solid curve depicts the flow synergies that accrue to (good) conglomerates as a function of segment distance. In the stationary version of the model, z_L and z_H are such that these flow synergies yield exactly the same payoff, which is intuitive: it does not matter whether synergies of a certain level S come from a segment combination that is above or below the ideal one, S is a sufficient statistic for the decision. In the dynamic model this is not the case, because firms know that on average they will remain a conglomerate for a number of years, and throughout that period the ideal segment distance is decreasing. As a rational response to this anticipated decrease, firms have an asymmetric merger-acceptance region, since they are willing to accept slightly lower payoffs today in exchange for a segment distance that will be closer to the (static) optimum in the future.

3 Calibration

So far we have shown data documenting the decline of corporate diversification and developed a model that can explain this pattern. In this section we ask whether the model can account for empirical patterns *quantitatively*, in the context of a calibration exercise.

Our strategy for the calibration has two main steps. First we take a steady-state version of the model (where σ is constant) and calibrate it to several corporate-diversification moments in data. Second, we use the parameters obtained from the first step to calibrate a model with time-varying σ, and test whether the model captures the trends we observe in data.

3.1 Steady-state approach

The steady-state model has two advantages: (i) given its tractability, the computational procedure for matching moments is relatively fast; (ii) there are no degrees of freedom associated with initial conditions (e.g., the initial proportion of single-segment firms). Naturally the steady-state model is inadequate to provide implications about how changes in technological specialization (σ) affect corporate-diversification *trends*,[12] but it provides a useful starting point. Furthermore, one would not expect specialization to be moving at a very fast pace, so the steady-state calibration should provide for a good approximation in terms of *levels*.

There are a total of seven parameters to calibrate: r (discount rate), λ_0 (likelihood of merger matches), λ_1 (likelihood of becoming bad conglomerate), λ_2 (likelihood of refocusing for bad conglomerates), β (overhead costs of bad conglomerates), ϕ (cost of project technological mismatch), and σ (inverse of technological specialization). A subset of the parameters are calibrated directly, namely r and σ. We set the discount rate r at 10%, which seems reasonable for a representative investor. We set $\sigma = 0.2$, which is just a normalization. As explained in section 2.1, in our model it is not straightforward to separately identify σ from ϕ, and such identification would hinge on particular assumptions about functional forms.

We use six moments in data as targets for calibrating the remaining five parameters. We describe the rationale for each choice below:

1. **Proportion of Conglomerates.** Our data counterpart to $1-p$, the fraction of diver-

[12] The only alternative would be a comparative-statics exercise, which would not consider that firms anticipate changes in σ.

sified firms in the economy, is the in-sample average proportion of book assets owned by conglomerates for the period 1998-2013, approximately 59% for the NAICS/I-O industry classification. We define this moment as the key one to match in the calibration (see details below). We focus on post-1997 data, given the change in segment reporting requirements and associated discontinuity in the proportion of single-segment assets in the economy (see bottom panels of figure 1 and related text). We focus on the NAICS/I-O classification given our later analysis of relatedness (see section 4.1).

2. **Single-Segment Value.** In data, the average market-to-book ratio of single-segment firms is 3.1, for the period 1990-2013, and it seems reasonable to assume that it embeds an expected growth rate of 2%. In our steady-state model there is no growth and thus we need to adjust the market-to-book ratio target accordingly. With a discount rate of 10%, the market-to-book ratio adjusted for no growth equals 2.5, which is therefore our target for J_0 (normalized single-segment value).

3. **Diversification Discount.** We compute the diversification discount in data by following the literature (see in particular Custódio, 2014). First we construct *excess value*, which is the log-difference between the market-to-book ratio of the firm (diversified or not) and a comparable portfolio of single-segment entities (see section A.5 in the appendix for more details). Then we run a regression of excess value on a constant and a diversification dummy. The negative of the coefficient on that dummy variable is usually interpreted as the diversification discount, and in our data it is 3.3%.[13] We match the model to this magnitude, where the theoretical diversification discount is computed using equations (8)-(10) and (14):

$$\frac{J_0 - (w\mathrm{E}[J_2] + (1-w)\mathrm{E}[J_1])}{J_0}.$$

4. **Probability of M&A.** Since most diversification is implemented via M&A, we would like the model to be realistic in terms of merger frequencies. The likelihood that a firm is involved in a takeover is 6% per year (Edmans, Goldstein, and Jiang, 2012), although this refers to any merger (including horizontal). Given the inclusion of non-diversifying

[13] This coefficient does not change much if we add control variables to the regression, although the inclusion of controls does reduce statistical significance.

mergers, one could be led to interpret the 6% as an upper bound. But on the other hand, the 6% figure includes single-segment firms and conglomerates, and the latter, according to our model, do not even engage in M&A. A further complication arises from the fact that in reality not all corporate-diversification moves occur via M&A, while in the model they do. In the end, we still target 6% as the rate of diversifying-merger activity for single-segment firms. While the comparison of model and data is not obvious for this moment, we believe the 6% figure is still informative about the general order of magnitude for the likelihood of diversifying. In the model, this likelihood corresponds to 1 minus the probability that the firm does not engage in any merger, which is given by

$$\sum_{k=0}^{\infty} \Pr\{matches = k\}(1-q)^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda_0}\lambda_0^k(1-q)^k}{k!} =$$
$$\frac{e^{-\lambda_0}}{e^{-\lambda_0(1-q)}} \underbrace{\sum_{k=0}^{\infty} \frac{e^{-\lambda_0(1-q)}[\lambda_0(1-q)]^k}{k!}}_{=1} = e^{-q\lambda_0}.$$

5. **Average Announcement Returns of Diversifying Mergers.** We attempt to match the average magnitude of diversifying-merger announcement returns, which in the model is simply
$$\frac{\mathrm{E}[J_1]-J_0}{J_0}.$$
In data, we use results from Akbulut and Matsusaka (2010), who report combined acquirer-target returns of 3.8% for cash deals. We focus on cash deals since we believe these are less influenced by signaling concerns (which we do not model).

6. **Refocusing Rate.** Finally, we attempt to match the rate of refocusing activity. In data, the average fraction of conglomerates becoming single-segment firms over a one-year period is 7.6%, for the period 1990-2013. The model counterpart to this magnitude is given by
$$w(1-e^{-\lambda_2}).$$

The procedure we use for generating parameters is as follows: (i) we require that the model matches precisely the fraction of assets allocated to conglomerates, since this is the

Table 1: Calibrated Parameters. The table shows the magnitude of each parameter used in the steady-state model calibration.

Description	Parameter	Value
Discount rate	r	0.10
Likelihood of merger matches	λ_0	0.31
Likelihood of becoming bad conglomerate	λ_1	0.07
Likelihood of refocusing bad congs.	λ_2	0.25
Overhead cost of bad conglomerates	β	0.39
Cost of project technological mismatch	ϕ	7.20
Inverse of technological specialization	σ	0.20

focus of our paper; (ii) regarding the other five moments, we minimize the equally-weighted sum of squared (relative) differences between model and data:

$$\sum_{i=1}^{5} \frac{1}{5} \left[\frac{target_i - output_i}{target_i} \right]^2$$

Table 1 summarizes the choice of parameters, and we note that assumption 1 is verified:

$$\pi_1(z^*(\phi,\sigma);\phi,\sigma) - \pi_0(\phi,\sigma) < \beta \Leftrightarrow 0.10 < 0.39$$

Table 2 reports key moments. The calibration yields a reasonable fit to data. The dimension which the model fits less well is the refocusing rate, where the calibration misses the actual magnitude by approximately 35%.

An interesting application of the calibration is that we can quantify the value of diversification, i.e., the value of all future diversification options. This is done by computing the difference between the actual value of single-segment firms (J_0), and their counterfactual value if diversification was not possible (π_0/r). Specifically, according to the model, we have that

$$\frac{J_0 - \pi_0/r}{J_0} = 3.1\%$$

of single-segment value can be attributed to diversification options. What is interesting about this result is that the value of diversification options is of the same order of magnitude as the diversification discount (in our data). This suggests that not adjusting for the value

Table 2: Model Outputs and Data. The table shows key moments, both in the calibration and in data; for the steady-state calibration. "Prop. Congs." is the proportion of assets in the economy allocated to diversified firms; "Single-Seg. Value" is the market-to-book ratio of single-segment firms; "Div. Discount" is the average valuation difference between a conglomerate and a comparable portfolio of specialized firms; "Probab. of M&A" stands for the likelihood that a single-segment BU engaged in at least one merger deal over a one-year period; "Av. Div. Returns" stands for the average announcement returns of diversifying mergers; and "Refocusing Rate" refers to the fraction of conglomerates becoming single-segment firms over a one-year period.

Moment	Model Counterpart	Calibration Output	Data/target
Prop. Congs.	$1-p$	59%	59%
Single-Seg. Value	J_0	2.9	2.5
Div. Discount	$\frac{J_0-(w\mathrm{E}[J_2]+(1-w)\mathrm{E}[J_1])}{J_0}$	3.4%	3.3%
Probab. of M&A	$1-e^{-\lambda_0 q}$	7.5%	6.0%
Av. Div. Returns	$\frac{\mathrm{E}[J_1]-J_0}{J_0}$	4.1%	3.8%
Refocusing Rate	$w(1-e^{-\lambda_2})$	4.8%	7.6%

of these options, as is normally done in the literature, may importantly over-state the level of the true diversification discount.

3.2 Time-varying technological specialization

This section builds on the steady-state calibration, adding a time-varying σ. Our objective is to compare the model output with the decline in corporate-diversification activity presented in section 1. The details of how the non-stationary model is solved are relegated to the appendix. In particular, we have to deal with the issue of having additional degrees of freedom associated with the choice of initial conditions, but such discussion detracts from economic intuition and thus is omitted from the main text. A summarized way to describe the procedure we implement is to view it as a choice of the rate at which σ decreases over time. We set the rate of growth of σ at -0.8%, in order to match a reasonable output growth rate in the economy. More specifically, our choice implies that single-segment firms' output increases at approximately 2% p.a. for the relevant time period. We also show in the appendix that the levels from the steady-state calibration (table 2) do not change significantly within the non-stationary model (table A.5), except for valuations which are

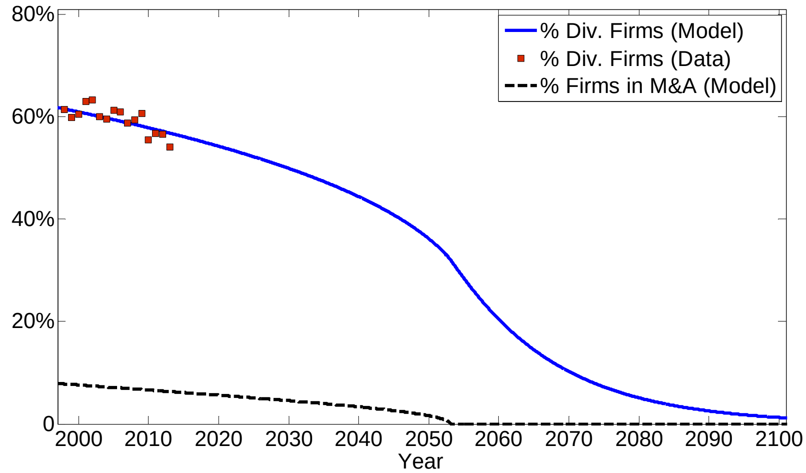

Figure 7: Evolution of Corporate Diversification: Model and Data. The figure shows three time series: (i) the fraction of conglomerate assets in data (square markers) for the period 1997-2013 and for the NAICS/I-O industry classification; (ii) the fraction of conglomerate assets in the model (solid line); and (iii) the (model) fraction of firms involved in diversifying mergers, in annualized terms (dashed line).

higher as intended (see discussion about target for J_0 in the previous section).

Figure 7 plots three time series: (i) the fraction of conglomerate assets in data (square markers) for the period subsequent to the change in reporting requirements (1997-2013), using the NAICS/I-O industry classification (same as third-row left panel in figure 1); (ii) the fraction of conglomerate assets in the model (solid line); and (iii) the fraction of firms involved in diversifying mergers (dashed line), in annualized terms, according to the model. The model fits the slope in data fairly well: in data, the fraction of conglomerate assets decreases at an average annual rate of 0.42 percentage points (p.pts.) from 1998 through 2013, with a model counterpart of 0.31 p.pts. Alternatively, we can say that the model generates a slope that is about 75% of the slope observed in data. This result is additional evidence for the model's ability to explain corporate-diversification patterns, especially given that this particular moment (i.e., the slope in the fraction of single-segment assets in the economy) had no direct bearing in the choice of parameters.

The model implies that an ever-growing technological specialization leads to a continued decline in corporate-diversification activity, with conglomerates (defined according to the 1997 NAICS/I-O industry classification) representing around 1% of the total assets in the economy by the end of the century. The model predicts an acceleration in the rate of conglomerate decline in the coming decades, driven by the vanishing of diversifying mergers—

the only corporate-diversification mechanism in this simplified economy—by the early 2050's.

A key driver for the decline of corporate-diversification activity in the model is the fact that the dark side of diversification is assumed to remain constant over time. Specifically, there is no change in the rate at which good conglomerates turn bad (λ_1), and there is no change in the extra overhead costs of bad conglomerates (β). Given these constant costs of diversification, coupled together with reduced benefits due to increased technological specialization, it is straightforward to see why the model predicts the demise of conglomerates. Although the assumption of constant diversification costs is somewhat extreme, one can alternatively interpret this as a normalization.[14] Under such interpretation, we could say that our calibration exercise strongly suggests that the benefits of corporate diversification decrease at a high rate, *relative* to the rate at which corporate-diversification costs decrease.

4 Testing additional implications of the model

In the previous section we developed a model that accounts for the observed decline in diversified firms. The model makes predictions along other dimensions than just the pervasiveness of conglomerates in the economy, and the objective of this section is to test whether these implications are borne out in data. First, we analyze the model's time-series and cross-sectional predictions about segment distance (section 4.1). Second, we turn to the time-series behavior of single-segment firm value and diversification discount (section 4.2).

4.1 Relatedness implications

The level of *relatedness* across segments has been a key variable in the study of conglomerates (Berger and Ofek, 1995; Fan and Lang, 2000), mostly to understand whether more-related conglomerates exhibit higher value. According to the mainstream view on relatedness, conglomerates that operate in unrelated industries are less focused and thus less profitable. Given the focus of our model on cross-division distance, it makes sense to investigate whether the model can account for empirical relatedness patterns, and that is the focus of this section.

Unfortunately, existing empirical relatedness measures do not have a natural spatial

[14] Even if we were to explicitly model time-varying benefits and costs, it is not clear that we could separately identify each slope.

interpretation, and therefore it is not intuitive to map these measures to the notion of segment distance in our model. We thus propose a novel spatial measure of relatedness, that takes into account the overall inter-industry architecture of the economy (section 4.1.1). Then we compare the empirical pattern with predictions from the calibrated model (sections 4.1.2-4.1.4)

4.1.1 A new empirical measure of relatedness

One of the contributions of our paper is a novel empirical measure of (un)relatedness, which in keeping with the model we also term segment distance. We compute segment distance in three steps: first we construct an economy-wide inter-industry network, using data from input-output tables; second, for all pairs of industries in the economy, we calculate how far they are located within the inter-industry network; and finally, for a particular conglomerate, we identify all relevant industry pairs and compute their average distance.[15] Our empirical relatedness measure maps naturally to the notion of segment distance in the theory model, and its key advantage (over traditional measures of relatedness) is that it takes the overall architecture of the economy into account. To illustrate the rationale for the latter claim, consider an example where two industries are vertically disconnected, but embedded in a broader supply chain (say because they share many customer and/or supplier industries). In such case it makes sense to say that these two industries are related, since they operate in the same (or similar) business environment.[16]

Next we turn to the details of how we empirically construct the segment-distance variable. The first step is to construct the inter-industry network, and we adopt the approach in Anjos and Fracassi (2015), who use the 1997 benchmark input-output tables at the detailed level. Focusing on just one year makes network measures immune to changes in industry classification, which is important for comparing segment distance over time.[17] Using the

[15] Fan and Lang (2000) also propose relatedness measures based on input-output flows, but do not consider the overall network architecture as we do.

[16] We note that the concept of "technology" in our model is quite broad, as in standard macroeconomic approaches, and includes a firm's managerial/organizational technology. Such technology is potentially similar for industries that are close-by in the economy-wide supply chain, even if the specific industry-level products are distinct.

[17] To illustrate the importance of reclassification at the detailed level, we note that there are 409 industries in 2002, versus 470 in 1997. Other recent papers building inter-industry networks from input-output tables focus on 1997 as well (Ahern and Harford, 2014; Anjos and Fracassi, 2015).

flows from the use tables for each industry pair (i, j), denoted as f_{ij},[18] we create a square matrix of normalized flows:

$$\overline{f}_{i,j} := \frac{0.5\,(f_{ij} + f_{ji})}{0.25\left(\sum_i f_{ij} + \sum_j f_{ij} + \sum_i f_{ji} + \sum_j f_{ji}\right)}.$$

This operation generates a symmetric square matrix of flows across industries. We employ a symmetric approach for simplicity and also because there is no clear way of assigning direction. Next we define an adjacent distance measure for each industry pair, d_{ij}, by taking the inverse of the normalized flow: $d_{ij} = 1/\overline{f}_{ij}$.

The second step is to construct an industry network, (a weighted undirected graph), using the adjacent distances d_{ij}. Given the industry network, we compute the weighted shortest path (one can think of distance as a cost) between any two industries, denoted as l_{ij}, by determining the total distance of the optimal path (i.e., the one that minimizes total distance or cost).[19]

The final step in the construction of segment distance is to average the measure across conglomerate industry pairs. Formally, conglomerate segment distance is then defined as

$$Seg.Dist. = \frac{\sum_{i \in \mathcal{I}} \sum_{j > i \wedge i \in \mathcal{I}} l_{ij}}{M(M-1)/2}, \tag{16}$$

where $\mathcal{I}$ denotes the set of industries a diversified firm participates in and M is the size of this set.

4.1.2 Time-series analysis

Figure 8 shows the evolution of segment distance, both in model and in data. The data spans the period 1990-2013. We start in 1990 because NAICS codes start in 1990 and are required to compute each diversified firm's industry portfolio in terms of 1997 Input-Output industries. Since the absolute level of segment distance does not have a meaningful

[18] Flows from the use tables report a dollar flow from commodity i to industry j, and where each industry has an assigned primary commodity.

[19] These network measures were computed using MATLAB BGL routines (available at `http://www.mathworks.com/matlabcentral/fileexchange/10922`), namely the dijkstra algorithm for minimal travel costs.

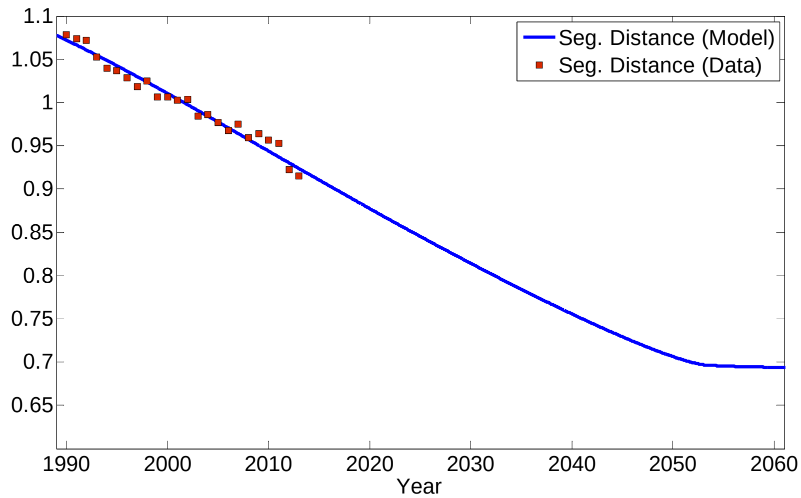

Figure 8: Evolution of Relatedness: Model and Data. The figure shows the evolution of segment distance in the model (solid line) and in data (square markers). Segment distance in data is computed as the average industry distance for all possible industry pairs where a conglomerate operates; industry distance is computed in the context of the inter-industry network induced by the 1997 benchmark input-output tables, following Anjos and Fracassi (2015). Segment distance in the model is the cross-sectional average of z (see section 2 for details).

interpretation (both in model and in data), we normalize segment distance by its average for the period 1990-2013; this makes the visual comparison of model and data easier.

The empirical evolution of segment distance is quite uncontroversial: for the period 1990-2013 there was a gradual, almost linear decrease in segment distance. The average yearly decline is about -0.7%, and the model matches this slope almost perfectly, despite not having been calibrated to fit this particular magnitude. This result is further evidence that our model does well in explaining corporate-diversification patterns.

Figure 8 also shows that the model predicts further declines in segment distance—or increases in relatedness—over the next decades, but there is a stabilization in the early 2050's. This stabilization occurs because diversifying mergers cease to take place at this time and there is henceforth no entry of fresh conglomerates with low average segment distance. When diversifying mergers cease, segment-distance dynamics are thus driven only by what happens to existing conglomerates. The slight downward slope in average segment distance occurs because bad conglomerates (i.e., those that incur overhead β and are on track for refocusing) are on average older and thus have higher average segment distance (reflecting "older" optimal policies); as these older conglomerates exit (i.e., refocus), the average segment distance of the remaining pool of conglomerates goes down.

Finally, in the appendix (section A.3) we also document time trends for two other relatedness measures that have been proposed in the literature, and all measures consistently exhibit an increase over time. This is evidence that, at least qualitatively, our results are not driven by the specific relatedness measure that we focus on.

4.1.3 Cross-sectional analysis

In previous sections we have focused on the time-series implications of our model. The model also has cross-sectional implications. Specifically, conglomerates should prefer intermediate segment distances, so as to optimize the returns to within-firm resource reallocation. Recalling the results from section 2.1 (see figure 5), too-low segment distance makes project swapping very frequent but with low reallocation gains per swap, whereas too-high segment distance implies very few reallocation opportunities. In this section we investigate these cross-sectional predictions.

The top-left panel of figure 9 describes the segment-distance distribution for our whole NAICS/I-0 sample, covering the period 1990-2013, and where we observe conglomerates cluster at intermediate distances. This is qualitatively consistent with our model, as visible in the simulated cross-sectional conglomerate distribution (for the comparable time period). Interestingly, although in the model we assume exogenous matches occur uniformly in terms of distance, the endogenous firm distribution is somewhat bell-shaped as well. This obtains because of the dynamics, for two reasons.[20] First, there is a reduction in the optimal segment distance over time, and thus, although there are legacy conglomerates at relatively high distances, most diversified firms cluster in the middle. Second, earlier conglomerates that merged when σ was relatively high were less stringent in terms of merger policies, i.e., the interval $[z_L, z_H]$ was wider (see figure 6). Therefore you also see some older legacy conglomerates at relatively low distances.

The top-right panel of figure 9 shows the empirical association between segment distance and conglomerate valuation. Here we should also observe a non-monotonic relationship—as illustrated by the simulation in the bottom-right panel—but the relationship is linear

[20] A caveat is in order. In theory, you could have *good* conglomerates that diversified a long time ago at the fringes of the merger-acceptance region and who, as σ reduces, could prefer to refocus. However, for simplicity, we assume that only bad conglomerates can refocus.

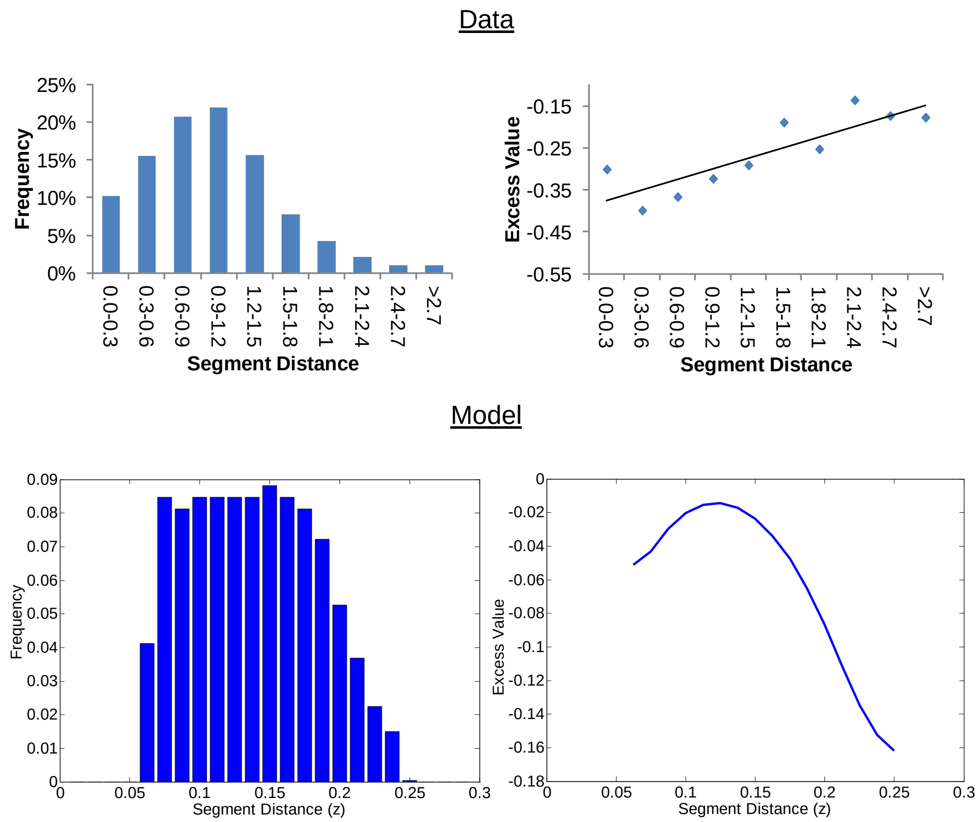

Figure 9: Cross-Sectional Analysis of Relatedness: Model and Data. The top-left panel shows the empirical segment-distance distribution for the period 1990-2013. The top-right panel shows conglomerate (empirical) average excess value, conditional on segment-distance class. Excess value is defined as the log-difference between the market-to-book ratio of a conglomerate and the market-to-book ratio 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. The bottom panels are simulations from the model for the equivalent time period.

and positive. In section 4.1.4 we address this mismatch between theory and data. We also find that the positive association between segment distance and excess value is robust to controlling for many other factors, as shown in table 3 (see summary statistics in section A.5).

Specification (1) presents the correlation between segment distance and excess value, but now controlling for year fixed effects, to account for macroeconomic shocks. 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), that are

Table 3: Excess Value and Segment Distance. The dependent variable is Excess Value, defined as the log-difference between the market-to-book ratio of a conglomerate and the market-to-book ratio of a similar portfolio of single-segment firms, following Berger and Ofek (1995). The table presents ordinary least squares regression coefficients, beta coefficients, and robust t-statistics clustered at the conglomerate level. The main explanatory variable is Segment Distance, defined as the average distance for every possible pair of industries that the conglomerate participates in, using the 1997 Input-Output industry classification system at the detailed level. All variables are defined in detail in the appendix. A constant is included in each specification but not reported in the table. All variables have been standardized. 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	0.087***	0.058***	0.053***	0.085***
	0.068	0.045	0.042	0.067
	(4.98)	(2.90)	(2.69)	(3.14)
Number of Segments		-0.040***	-0.053***	-0.069***
		-0.052	-0.070	-0.091
		(-3.83)	(-4.78)	(-5.38)
Weight rel. segs.		0.190***	0.165***	0.120**
		0.053	0.046	0.034
		(3.50)	(3.10)	(2.00)
Vert. rel.		-0.002	-0.005	0.018**
		-0.007	-0.017	0.069
		(-0.55)	(-1.23)	(2.44)
Excess centrality		0.337***	0.322***	0.494***
		0.056	0.054	0.082
		(3.62)	(3.48)	(3.52)
Excess assets			0.022***	0.022*
			0.073	0.071
			(4.02)	(1.95)
Excess EBIT/sales			-0.004***	-0.002***
			-0.102	-0.059
			(-6.60)	(-5.03)
Excess capex/sales			-0.001	-0.001
			-0.005	-0.010
			(-0.79)	(-1.59)
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	No	Yes
R^2	0.023	0.028	0.043	0.040
Obs.	29,221	29,221	28,058	28,058

traditionally associated with the level of business focus. It also includes a vertical-relatedness measure, computed following Fan and Lang (2000), which allows us to differentiate the effects of segment distance 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 (2015), 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. Specification (3) adds financial variables to the regression, constructed according to the the approach recommended in Gormley and Matsa (2013),[21] and specification (4) includes firm fixed effects.

Segment distance has an economically-significant impact in terms of conglomerate value, as shown by its beta coefficient. A one-standard-deviation increase in segment distance is associated with an increase of between 0.042 and 0.068 standard deviations in excess value. Excess value has a standard deviation of 0.72, so this corresponds to an increase of between 0.03 and 0.05 in excess value. Since the average conglomerate has an excess value that is 32% lower than its benchmark portfolio, then this effect is between $0.03/0.68 \approx 4.4\%$ and $0.05/0.68 \approx 7.4\%$ of firm value for the average conglomerate.

In table 3, and in the most stringent regression, the coefficients on number of segments, weight of related segments, and vertical relatedness are all consistent with previous literature: relatedness and focus is associated with higher firm value.[22] 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 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 suggests 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

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

[22] We note however that the number of segments negative coefficient can be picking up costs of organizational complexity and is not necessarily driven by relatedness (see Sanzhar, 2006 for more details).

non-redundant information (excess centrality).

4.1.4 Reconciling model and cross-sectional evidence

A possible explanation for the linear (instead of non-monotonic) relationship between segment distance and excess value would be that merger opportunities take place only in a relatively close neighborhood of the firm's core activities. There are plausible reasons for this "home bias", for example adverse selection being more of a concern for distant mergers. The initially positive association between segment distance and frequency, shown in the top-left panel of figure 9, 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 market-to-book ratios for high-segment-distance firms). We also note that there is evidence in other settings that firms are more likely to engage in localized M&A activity, both geographically and culturally (Ahern, Daminelli, and Fracassi, 2012).

Whereas the explicit modeling of informational frictions is outside the scope of our paper, it is straightforward to change which merger matches occur, and in particular we can require that they take place within a neighborhood of the firm's business environment. In the appendix (section A.4) we present an extension of our main model where matches are truncated. We calibrate this model to data and show that the extended model can accommodate the positive association between segment distance and excess value.

4.2 Time-series value implications

Our final analysis focuses on the model's value implications. The top-left panel of figure 10 shows the evolution of the goodwill-adjusted market-to-book ratio for single-segment firms, for the period 1990-2013 (NAICS/I-O classification). The market-to-book ratio for the average single-segment firm went up by about 0.06 points per year. On the other hand, there is no consistent change in the diversification discount, which we compute year-by-year in the top-right panel of figure 10.

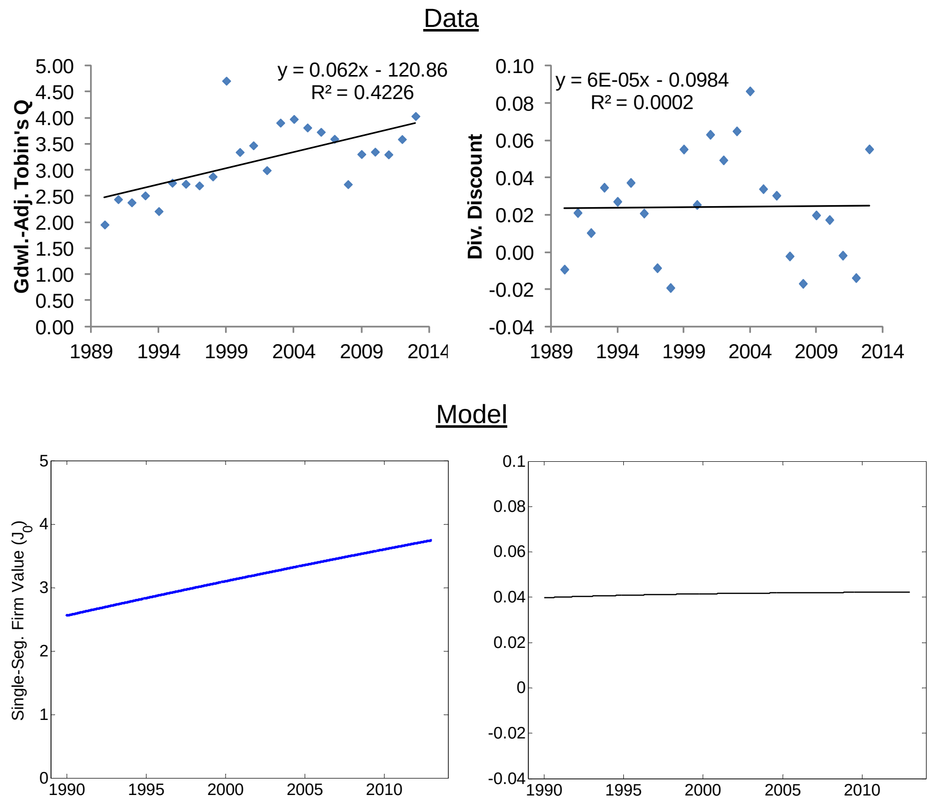

Figure 10: Evolution of Firm Value: Model and Data. The top-left panel shows the evolution of cross-sectional-average market-to-book ratio (adjusted by goodwill, following Custódio, 2014), for single-segment firms (diamond markers) and conglomerates (square markers), using the NAICS/I-O industry classification. The top-right panel shows the evolution of the diversification discount, computed in two steps. First, following Berger and Ofek (1995), we construct excess value, the log-difference between a firm's market-to-book ratio and the market-to-book ratio of an assets-weighted comparable portfolio of single-segment firms. Second, we run a year-by-year regression of excess value on a constant and a diversification dummy. The diversification discount is the negative of the dummy coefficient. The bottom panels show the model counterparts.

The bottom-left panel shows the calibration predictions for single-segment firm value, for the equivalent time period. The (normalized) value of single-segment firms increases at a rate similar to data (0.05 points per year), although the model was not calibrated to fit this particular magnitude. Our calibration also matches the trend of the diversification discount. This result is interesting, because one can pick alternative parameters—rather than the ones in the calibration—where the model implies that the diversification discount is changing at a significant rate over time. In that sense, we find that the calibration seems consistent with an additional moment that it is not explicitly designed to match. To understand this better, recall that, in the model, the diversification discount is

$$\frac{J_0 - E[J_{div}]}{J_0},$$

where $E[J_{div}]$ is the average market-to-book ratio of conglomerates. The numerator in the above fraction is increasing over time, because single-segment firms benefit more from technological specialization than conglomerates do. However, the denominator is also increasing. Therefore, in our model, depending on parameter choice, the diversification discount can trend upwards or downwards over time.

5 Conclusion

We document that conglomerates have been declining since the 1970's, and, moreover, that these firms' divisions have become more related over time. We develop a simple dynamic trade-off model of corporate diversification that can account for these patterns, as well as other empirical features of corporate diversification. In the model, diversification synergies stem from the ability to undertake efficient within-firm (technological) resource reallocation (bright side). These benefits are traded off with the possibility that the conglomerate incurs organizational-complexity costs (dark side). Over time, technological specialization increases, i.e., business units automatically focus more on what they do best, and the need/opportunity to reallocate resources gradually disappears. Therefore, synergies from diversification decrease as well, and conglomerates tend to decline in equilibrium. While other potential explanations exist for the decline of conglomerates—e.g., change in the benefits of coinsurance or internal capital markets—, our explanation is parsimonious and consistent with data. In short, we contribute to the literature by providing new evidence about corporate diversification, together with a simple model that accounts reasonably well for observed empirical patterns.

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2012, The network origins of aggregate fluctuations, *Econometrica* 80, 1977–2016.

Ahern, Kenneth R., Daniele Daminelli, and Cesare Fracassi, 2012, Lost in translation? The effect of cultural values on mergers around the world, *Journal of Financial Economics* (forthcoming).

Ahern, Kenneth R., and Jarrad Harford, 2014, The importance of industry links in merger waves, *Journal of Finance* 69, 527–576.

Akbulut, Mehmet E., and John G. Matsusaka, 2010, 50+ years of diversification announcements, *Financial Review* 45, 231–262.

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, 2015, Shopping for information? Diversification and the network of industries, *Management Science* 61, 161–183.

Becker, Gary S., and Kevin M. Murphy, 1994, *The division of labor, coordination costs, and knowledge*, chapter *in* Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition) (The University of Chicago Press).

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.

Coase, Ronald, 1937, The nature of the firm, *Economica* 4, 386–405.

Comment, R, and G. H. Jarrell, 1995, Corporate focus and stock returns, *Journal of Financial Economics* 37, 67–87.

Custódio, Cláudia, 2014, Mergers and acquisitions accounting and the diversification discount, *Journal of Finance* 69, 219–240.

David, Joel, 2014, The aggregate implications of mergers and acquisitions, Working paper, available at SSRN.

Denis, D. J., D.K. Denis, and A. Sarin, 1997, Agency problems, equity ownership, and corporate diversification, *Journal of Finance* 52, 135–160.

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

Dimopoulos, Theodosios, and Stefano Sacchetto, 2014, Merger activity in industry equilibrium, Working paper, available at SSRN.

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).

Giroud, Xavier, and Holger M. Mueller, 2015, Capital and labor reallocation within firms, *Journal of Finance* (forthcoming).

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* 27, 617–661.

Graham, John. R., Michael L. Lemmon, and Jack G. Wolf, 2002, Does corporate diversification destroy value?, *Journal of Finance* 57, 695–719.

Hackbarth, Dirk, and Erwan Morellec, 2008, Stock returns in mergers and acquisitions, *Journal of Finance* 63, 1213–1252.

Hubbard, R. Glenn, and Darius Palia, 1999, A reexamination of the conglomerate merger wave in the 1960s: An internal capital markets view, *Journal of Finance* 54, 1131–1152.

Jackson, Matthew O., 2008, *Social and Economic Networks* (Princeton University Press).

Kaplan, S., and M. S. Weisbach, 1992, The success of acquisitions: evidence from divestitures, *Journal of Finance* 47, 107–138.

Lang, Larry H. P., and René M. Stulz, 1994, Tobin's q, corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248–1280.

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.

Morellec, Erwan, and Alexei Zdhanov, 2008, Financing and takeovers, *Journal of Financial Economics* 87, 556–581.

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)).

Rosen, Sherwin, 1978, Substitution and division of labour, *Economica* 45, 235–250.

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.

Tate, Geoffrey, and Liu Yang, 2015, The bright side of corporate diversification: evidence from internal labor markets, *Review of Financial Studies* (forthcoming).

Varian, Hal R., 2010, Computer mediated transactions, *American Economic Review* 100, 1–10.

Williamson, Oliver, 1975, *Markets and Hierarchies: Analysis and Antitrust Implications* (Free Press).

Yang, Liu, 2008, The real determinants of asset sales, *Journal of Finance* 63, 2231–2262.

Yang, Xiaokai, and Jeff Borland, 1991, A microeconomic mechanism for economic growth, *Journal of Political Economy* 99, 460–482.

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

Appendix

TABLE OF CONTENTS

A.1 Proofs

Proof of proposition 1.

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

<u>Case 1:</u> $z \leq \sigma$

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 is the one that minimizes total "travel" from the (assigned) projects to each division/BU. Inspection of the different possibilities allows us to determine the optimal policy for each case, with results shown in table A.1.

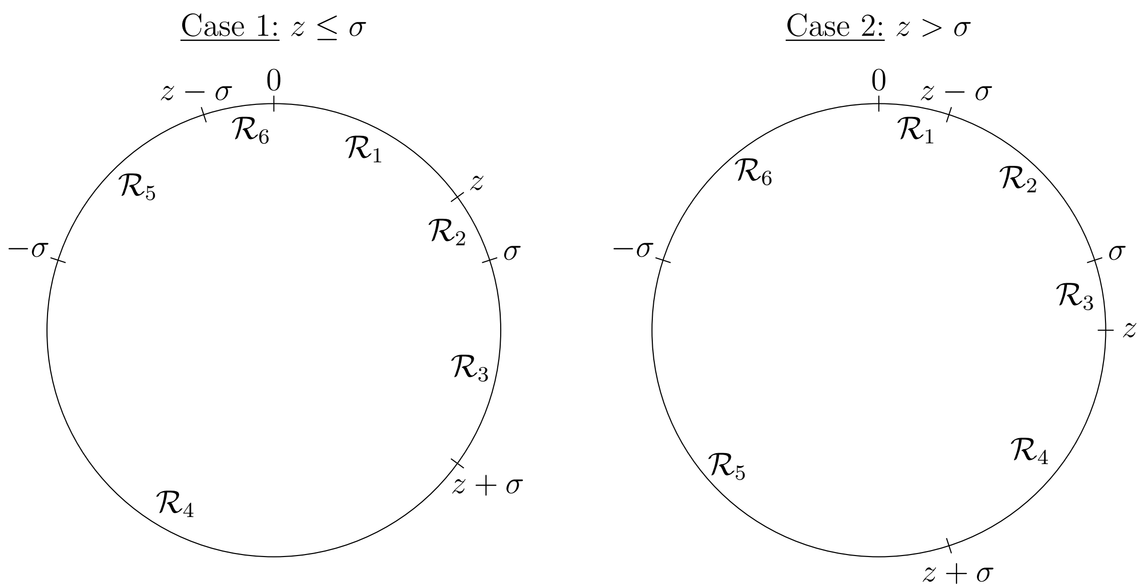

Figure A.1: Splitting the circle into regions. In the left example, $\sigma = 0.2$ and $z = 0.15$. In the right example, $\sigma = 0.2$ and $z = 0.25$.

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.1: Optimal allocation policy (swap/no-swap) 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.2.

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

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

We are omitting the explicit integration procedures, since all conditional distributions are uniform (in the relevant region), so probabilities and expected distances are generally simple functions of (region) arc length; the slightly more complex case is the computation of $\mathrm{E}[\min(z_{i,P_i}, z_{j,P_j})|...]$, 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.2 into equation (A.1), and after a few steps of algebra, one obtains

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

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

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.3: Optimal allocation policy (swap/no-swap) as a function of project location; with $z > \sigma$.

which implies equation (3a) in the proposition.

Case 2: $z > \sigma$

For this case let us make the additional assumption that $z \leq 2\sigma$. This assumption is made without loss of generality, since for $z > 2\sigma$ there cannot be any gains from diversification and the two-division conglomerate is simply a collection of two specialized business units, each undertaking its own projects (this corresponds to equation (3c) 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 regions 2, 3, or 4. Table A.3 shows the optimal allocation policy for each scenario.

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

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

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

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

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

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

which implies expression (3b) 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 (3a), 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 (4) 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 profit function. To prove this, it is sufficient to show that equation

(3b) 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.

We focus on the equilibrium where mergers take place (the other case is trivial). The solution to the firm's optimization problem (7) 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^* (where $\pi_1^G(z_L)=\pi_1^G(z_H)$). The solution to the problem described in expression (7), given financial markets' equilibrium, needs to verify the following conditions (where for notational simplicity we set $\tau=0$):

$$\begin{cases} rJ_2(z,t)\,dt=[\pi_1(z)-\beta]\;dt+\mathrm{E}_t[dJ_t] \\ rJ_1(z,t)\,dt=\pi_1(z)\,dt+\mathrm{E}_t[dJ_t] \\ rJ_0\,dt=\pi_0\,dt+\mathrm{E}_t[dJ_t] \end{cases}$$

Given the assumed Poisson processes and the conjectured merger-acceptance probability q, the above system can be written as

$$\begin{cases} rJ_2(z,t)=[\pi_1(z)-\beta]+\lambda_2[J_0-J_2(z)] & \text{(A.5)} \\ rJ_1(z,t)=\pi_1(z)+\lambda_1[J_2(z)-J_1(z)] & \text{(A.6)} \\ rJ_0=\pi_0+\lambda_0 q\left\{\mathrm{E}[J_1(z,t+dt)|z\in[\underline{z},\overline{z}]]-J_0\right\}. & \text{(A.7)} \end{cases}$$

Manipulation of equations (A.5)-(A.6) straightforwardly yields expressions (8)-(9) in the proposition. Using equation (9), we can write

$$\mathrm{E}[J_1(z,t+dt)|z\in[\underline{z},\overline{z}]]$$

as

$$\int_{z_L}^{z_H}\left(\frac{1}{z_H-z_L}\right)J_1(z)\,dz=\frac{\overline{\pi}_1(r+\lambda_1+\lambda_2)-\lambda_1\beta+\lambda_1\lambda_2 J_0}{(r+\lambda_1)(r+\lambda_2)}.$$

Inserting the above expression into equation (A.7), and solving for J_0, one obtains equation

(10) in the proposition.■

Proof of proposition 4.
Let us begin with the second result in the proposition. Since in equilibrium the distribution of firms is stationary, it needs to be the case that the mass of good conglomerates becoming bad over an infinitesimal dt, $(1-p)(1-w)\lambda_1\, dt$, be the same as the mass of bad conglomerates refocusing, which is $(1-p)w\lambda_2\, dt$. Simplification of this equality yields expression (14) in the proposition. The first result obtains along similar lines. The mass of single-segment firms becoming diversified over an infinitesimal dt, $p\lambda_0 q\, dt$, must be the same as the mass of firms refocusing, which is $(1-p)w\lambda_2\, dt$. Using the expression for w and simplifying yields equation (13). Next we turn to the third 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:

$$J_1(z^*) \le \frac{\pi_0}{r} \Leftrightarrow \frac{\pi_1(z^*)(r+\lambda_1+\lambda_2) - \lambda_1\beta + \lambda_1\lambda_2 J_0}{(r+\lambda_1)(r+\lambda_2)} \le \frac{\pi_0}{r},$$

where we used equation (9). Replacing π_0 and $\pi_1(z^*)$ by their expressions as a function of primitives σ and ϕ (equations (2) and (5)); and after a few steps of algebra, yields the result $\phi\sigma \le C$. For the necessity part of the proof we note that $q = 0$ could not be an equilibrium if $\phi\sigma > C$, since, by the argument above, there would be some mergers worth executing (which is inconsistent with $q = 0$).■

Proof of proposition 5.
First note that the equilibrium exists and is unique for $\phi\sigma \le C$, where C 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 $\phi\sigma > C$. Since $J_1(z^*) > J_0 > J_1(0)$, and given continuity, this implies that there exists $\{z_L, z_H\}$ such that $J_1(z_L) = J_1(z_H) = J_0$. Uniqueness follows from continuity and the fact that the equilibrium is unique at $\phi\sigma \le C$ (see for example Garcia and Zangwill, 1982 for more technical details).■

A.2 Details about calibration with time-varying σ

A.2.1 Solution method

With time-varying σ firms still face the optimization problem described in (7), except now merger-acceptance policies are time-varying. The only caveat to the similarity in optimization problems is that for the non-stationary model, some firms could actually merge at some point in time and refocus later, *even without turning into bad conglomerates.* This could take place for pairs that were close to the exercise boundary, and for whom the opportunity cost of not being in the mergers market increases over time (as σ decreases). This caveat notwithstanding, and for simplicity, we assume that only bad conglomerates can refocus.

We solve the model using the following steps:

1. We first determine a level of σ for date 0, and we choose a value that is high but still produces strictly positive average profits for single-segment firms (see equation (2)). Specifically, we set this magnitude at 0.26. Even for this higher value of σ, assumption 1 is verified:
$$\pi_1(z^*(\phi,\sigma);\phi,\sigma) - \pi_0(\phi,\sigma) < \beta \Leftrightarrow 0.13 < 0.39.$$

2. Using the starting value for σ we solve the steady-state model and obtain distributions for firm types, namely the proportion of single-segment firms p_0 and the initial fraction of bad conglomerates w_0. This procedure allows us to use initial conditions that are not excessively arbitrary.

3. We then choose a terminal time horizon T, which we pick to be 150 years, and a terminal level of σ (set at 0.07, which implies an output growth rate of 2% for single-segment firms, within the relevant time period of 1998-2013). We conjecture that the terminal level of σ is such that mergers no longer take place on the last period, which allows us to compute J_0 at the terminal date T simply as $\pi_{0,T}/r$ (implicitly we assume that σ is constant for periods later than T). This is an important input for the calculation of value functions in previous periods. Similarly, we can compute theoretical values for $J_1(z)$ and $J_2(z)$ at time T using equations (9) and (8).

4. We discretize time using an interval of length δ_t (1 week in our numerical implementation), and obtain each relevant value function, under the assumption of a particular

policy path $\{z_{L,t}, z_{H,t}\}_{t\in[0,T]}$. In particular, value functions are obtained recursively, by using the following system of finite differences (these basically discretize the non-stationary version of the differential equations presented in the proof of proposition 3):

$$\begin{aligned}
J_2(z,t) &= (\pi_1(z,t)-\beta)\delta_t + (1-r\delta_t)\left[\lambda_2\delta_t J_0(t+1) + (1-\lambda_2\delta_t)J_2(z,t+1)\right] \\
J_1(z,t) &= \pi_1(z,t)\delta_t + (1-r\delta_t)\left[\lambda_1\delta_t J_2(z,t+1) + (1-\lambda_1\delta_t)J_1(z,t+1)\right] \\
J_0(t) &= \pi_0(t)\delta_t + (1-r\delta_t)\left[\lambda_0\delta_t q(t+1)\mathrm{E}[J_1(z,t+1)|z\in[z_{L,t},z_{H,t}]]+\right. \\
&\quad \left.+(1-\lambda_0\delta_t q(t+1))J_0(t+1)\right]
\end{aligned}$$

5. We iterate the policy function using the optimal decision rule (i.e, merge only if it creates value), and obtain convergence.

6. Given the sequence of merger-acceptance policies, we compute the laws of motion for each mass of firm types; we denote the time-t density (at z) of bad conglomerates as $c_b(z,t)$ and the density of good conglomerates as $c_g(z,t)$:

$$\begin{aligned}
\frac{dp(t)}{dt} &= \int_0^{1/2} c_b(z,t-1)\lambda_2\, dF(z) - p(t-1)\lambda_0 q(t) \\
\frac{dc_g(z,t)}{dt} &= p(t-1)\lambda_0 \mathbb{1}_{z\in[z_{L,t},z_{H,t}]}\, dF(z) - c_g(z,t-1)\lambda_1 \\
\frac{dc_b(z,t)}{dt} &= c_g(z,t-1)\lambda_1 - c_b(z,t-1)\lambda_2,
\end{aligned}$$

where $F(z)$ is the uniform distribution with support $[0,1/2]$.

7. With the firm-type distributions and value functions it is straightforward to obtain all outputs. The relevant period is identified by finding the time step at which $\sigma = 0.2$ (the choice in the steady-state calibration) and determining that to be the midpoint of the 1998-2013 interval.

A.2.2 Additional outputs

Table A.5 shows that the magnitudes implied by the steady-state model in terms of levels are close to those generated by the non-stationary calibration.

Table A.5: Model Outputs and Data: Steady-State vs. Non-Stationary model. The table shows key moments, both in the steady-state (SS) calibration and the non-stationary (NS) calibration (averages across 1998-2013 period), as well as data/targets. "Prop. Congs." is the proportion of assets in the economy allocated to diversified firms; "Single-Seg. Value" is the market-to-book ratio of single-segment firms; "Div. Discount" is the average valuation difference between a conglomerate and a comparable portfolio of specialized firms; "Probab. of M&A" stands for the likelihood that a single-segment BU engaged in at least one merger deal over a one-year period; "Av. Div. Returns" stands for the average announcement returns of diversifying mergers; and "Refocusing Rate" refers to the fraction of conglomerates becoming single-segment firms over a one-year period.

Moment	SS-Calibration	NS-Calibration	Data/target
Prop. Congs.	59%	59%	59%
Single-Seg. Value	2.9	3.4	2.5 (SS) / 3.1 (NS)
Div. Discount	3.4%	4.2%	3.3%
Probab. of M&A	7.5%	7.1%	6.0%
Av. Div. Returns	4.1%	3.3%	3.8%
Refocusing Rate	4.8%	4.9%	7.6%

The main differences are a higher value of single-segment firms J_0 (as intended, see discussion in section 3), as well as lower average diversifying-merger returns (although calibrated announcement returns are 87% of the average return in data, which is a good fit). The higher J_0 is to be expected, since now value functions incorporate growth in cash flows. Furthermore, a higher J_0 makes returns to diversification lower.

A.2.3 Robustness check

This section presents a simple robustness check of our results, where we ask how much initial conditions matter. To address this issue we simulate the non-stationary model, but adopting rather extreme initial conditions, in particular that all firms are single-segments; and that all conglomerates are good.

Figure A.2 plots the evolution of p, the fraction of single-segment firms, for this new simulation; and compares this output with the output of our main non-stationary calibration. In particular, if one focuses on the relevant period 1990-2013, one observes little difference between the main simulation path and the alternative one. For the sake of space we do not report other magnitudes, but the differences are also small. The key takeaway of this analysis is that our results do not seem to be driven by our treatment of initial conditions, the effect of which vanishes relatively quickly.

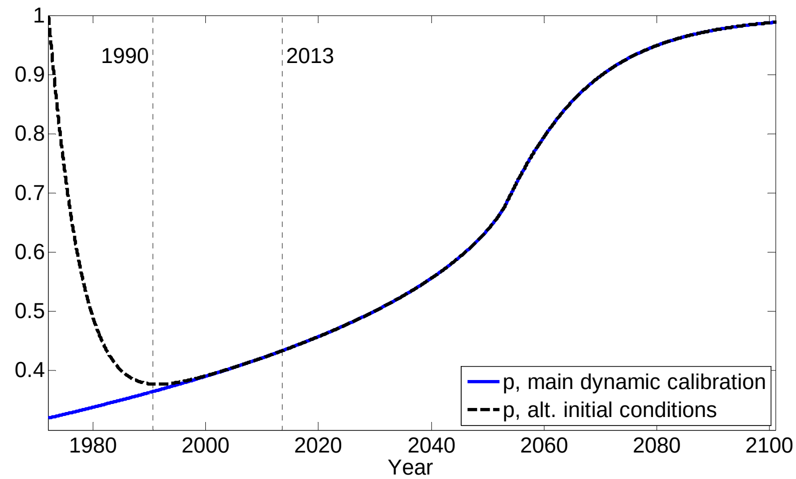

Figure A.2: Initial Conditions: Robustness Check. The figure plots the evolution of p under alternative initial conditions: 99.9% of all firms are single-segment at time 0; and 99.9% of all conglomerates are good at time 0.

A.3 Additional relatedness trends

Figure A.3 shows the evolution of alternative relatedness measures. The top panel shows the evolution of vertical relatedness, which, following Fan and Lang (2000) measures the intensity of the vertical connection between the primary-segment industry and the secondary-segment industries. The bottom panels show the evolution of the weight of related segments, as a fraction of the total number of segments. In the bottom-left (bottom-right) panel, related segments are the number of unique segments of a conglomerate using the detailed 1997 Input-Output industry classification system (4-digit SIC system), minus the number of unique segments of a conglomerate using the 3-digit 1997 Input-Output industry classification system (3-digit SIC system), following Berger and Ofek (1995). All three relatedness measures exhibit an increasing trend, in line with our main (un)relatedness measure (segment distance).

A.4 Extension: model with truncated matching

In this section we extend our model to allow for a truncation in the distribution of merger matches. In particular, we assume that matches only occur within a neighborhood of the firm's business environment. We define this truncation in the simplest possible way, requiring

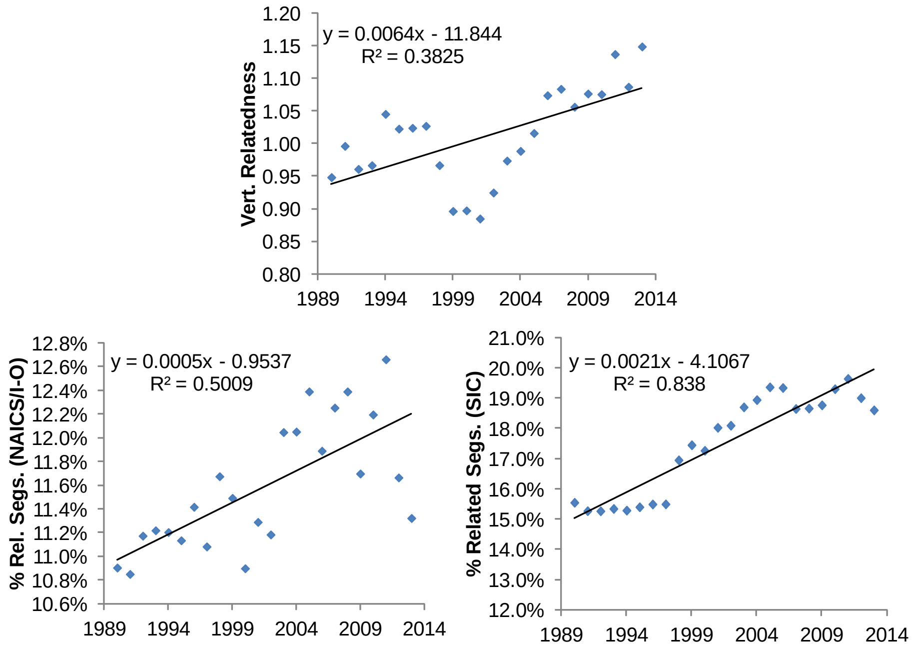

Figure A.3: Evolution of Relatedness: Alternative Empirical Measures. The top panel shows the evolution of vertical relatedness, computed following Fan and Lang (2000). The bottom-left (bottom-right) panel shows the evolution of the ratio between the number of related segments, based on the NAICS/I-0 (SIC) classification. All variables are defined in detail in section A.5.

that matches occur uniformly in the interval $[0, z_{max}]$. When this new constraint is binding, we are able to partly match the cross-sectional empirical pattern presented in section 4.1.3, namely that excess value increases in segment distance. Given the focus on the cross section, we calibrated the extended model only for the steady-state case. We expand the set of moments from the baseline case to include the difference in excess value across high- and low-segment-distance conglomerates,

$$\Delta EV := EV|_{z>median} - EV|_{z\leq median} \tag{A.8}$$

which in our data is about 0.087.

Table A.6 summarizes the choice of parameters, and assumption 1 is once more non-binding:

$$\pi_1(z^*(\phi,\sigma);\phi,\sigma) - \pi_0(\phi,\sigma) < \beta \Leftrightarrow 0.11 < 0.26$$

Table A.7 reports key levels (compares to table 2 for the main model). The truncated model

Table A.6: Calibrated Parameters: Baseline vs. Truncated model. The table shows the magnitude of each parameter, both in the baseline steady-state calibration and the truncated steady-state calibration.

Description	Parameter	Baseline	Truncated
Discount rate	r	0.10	0.10
Likelihood of merger matches	λ_0	0.31	0.15
Likelihood of becoming bad conglomerate	λ_1	0.07	0.06
Likelihood of refocusing bad congs.	λ_2	0.25	0.21
Overhead cost of bad conglomerates	β	0.39	0.26
Cost of project technological mismatch	ϕ	7.20	8.30
Inverse of technological specialization	σ	0.20	0.20
Maximal matching distance	z_{max}	n.a.	0.05

Table A.7: Model Outputs and Data: Baseline vs. Truncated model. The table shows key moments, both in the baseline steady-state calibration and the truncated steady-state calibration, as well as data/targets. "Prop. Congs." is the proportion of assets in the economy allocated to diversified firms; "Single-Seg. Value" is the market-to-book ratio of single-segment firms; "Div. Discount" is the average valuation difference between a conglomerate and a comparable portfolio of specialized firms; "Probab. of M&A" stands for the likelihood that a single-segment BU engaged in at least one merger deal over a one-year period; "Av. Div. Returns" stands for the average announcement returns of diversifying mergers; "Refocusing Rate" refers to the fraction of conglomerates becoming single-segment firms over a one-year period; and "Δ Excess Value" is the difference in excess value between above-median-segment-distance and below-median-segment-distance conglomerates.

Moment	Baseline	Truncated	Data/target
Prop. Congs.	59%	59%	59%
Single-Seg. Value	2.9	1.8	2.5
Div. Discount	3.4%	3.2%	3.3%
Probab. of M&A	7.5%	6.5%	6.0%
Av. Div. Returns	4.1%	6.4%	3.8%
Refocusing Rate	4.8%	4.2%	7.6%
Δ Excess Value	0.0%	5.4%	8.7%

can fit data well (better than the baseline model in some dimensions, worse in others), and in particular explains 60% of the cross-sectional relation between segment distance and excess value (Δ EV is 5.4% in the model and 8.7% in data). The main difference in the parameters we were already using in the main model is the choice of λ_0. In the truncated model, $\lambda_0 = 0.15$, whereas $\lambda_0 = 0.31$ in the main model. The difference is explained by the fact that in the main model, there are matches that occur beyond the useful range, i.e., at distances bigger than 2σ (unlike with truncated matching). Therefore, in order to obtain the same rate of merger activity, there need to be more matches taking place.

A.5 Summary statistics and variable definitions (NAICS/I-O)

- *Assets*: The total assets of a company (Source: AT variable in COMPUSTAT). This variable is adjusted for goodwill (Source: GDWL variable in COMPUSTAT), following Custódio (2014).

- *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 Segments 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 Segments 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 1997 Input-Output industry classification system (Source: COMPUSTAT, COMPUSTAT Segments, BEA, and Authors' Calculations). We scale the raw variable by its unconditional mean.

- *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 Segments and Authors' Calculations).

- *Excess Value*: The log-difference between the market-to-book ratio of a conglomerate (adjusted for goodwill following Custódio, 2014) and the assets-weighted market-to-book ratio of a similar portfolio of single-segment firms, using the detailed 1997 Input-Output industry classification system (Source: CRSP, COMPUSTAT, BEA, and Authors' Calculations).

- *Market-to-Book Ratio*: The sum of total goodwill-adjusted assets (AT-GDWL), following Custódio (2014), 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 adjusted by goodwill (AT-GDWL) (Source: COMPUSTAT).

- *Number of Segments*: The number of unique segments of a conglomerate using the detailed 1997 Input-Output industry classification system (Source: COMPUSTAT Segments and BEA).

- *Related Segments*: The number of unique segments of a conglomerate using the detailed 1997 Input-Output industry classification system, minus the number of unique segments of a conglomerate using the 3-digit 1997 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.

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

- *Weight Related Segments*: The ratio of Related Segments by Number of Segments (Source: Authors' Calculations).

Table A.8: Summary Statistics. The table presents summary statistics for each variable, comprising the period 1990-2013.

Panel A: Conglomerates

Variable	Mean	Std. Dev.	Min.	Max.	#Obs.
Market-to-Book Ratio	1.91	2.04	0.48	55.47	29,221
Excess Value	-0.32	0.72	-3.67	7.89	29,221
Segment Distance	1.00	0.56	0.03	4.40	29,221
Excess Centrality	1.00	0.12	0.85	2.16	29,221
Vert. Relatedness	1.00	2.70	0	24.80	29,221
N. Segments	2.61	0.94	2	10	29,221
Weight Rel. Segments	0.12	0.20	0	0.86	29,221
Assets	4,560	16,269	0.02	492,735	29,221
EBIT/sales	-0.14	8.72	-1,018	642.3	28,407
capex/sales	0.14	2.89	-0.94	433.1	28,881
Excess Assets	-0.10	2.37	-10.78	10.62	29,221
Excess EBIT/sales	3.67	18.70	-1,018	720.5	28,395
Excess capex/sales	-0.71	6.92	-282.5	433.0	28,870

Panel B: Single-Segment Firms

Variable	Mean	Std. Dev.	Min.	Max.	#Obs.
Market-to-Book Ratio	3.14	4.90	0.48	63.41	95,348
Excess Value	-0.28	0.70	-3.81	3.47	95,348
Assets	1,017	4,923	0.001	244,193	95,348
EBIT/sales	-7.35	175.9	-28,838	5,638	88,205
capex/sales	1.23	47.15	-17.92	7,826	94,027