

Technological Specialization and Corporate Diversification*

Fernando Anjos[†]

Cesare Fracassi[‡]

Abstract

We document a trend towards fewer and more-focused conglomerates, and develop a model that explains these patterns based on increasing technological specialization. In the model, diversification adds value by allowing efficient within-firm resource reallocation. However, synergies decrease with technological specialization, leading to fewer diversified firms over time. Also, the optimal level of technological diversity across conglomerate divisions decreases with technological specialization, leading to more-focused conglomerates. The calibrated model matches the evolution of conglomerate pervasiveness and focus, and other empirical magnitudes: growing output, level and trend of the diversification discount, frequency and returns of diversifying mergers, and frequency of refocusing activity.

April 21, 2014

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 Altı, Matt Rhodes-Kropf (AFA discussant), Alessio Saretto, Laura Starks, and Malcolm Wardlaw. The authors also thank comments from seminar participants at the University of Texas at Austin, and participants at the following conferences: 2012 European meetings of the Econometric Society, 2013 North American Summer meetings of the Econometric Society, and 2014 meetings of the American Finance Association.

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

1 Introduction

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

If technological specialization is ever-increasing, one would expect conglomerates to also become more focused, or less diverse, over time. This is indeed what we find in data, using an input-output-based measure of technological diversity: In the last two decades, technological diversity across divisions decreased approximately 12% for the average conglomerate. We also document an increase in the fraction of assets allocated to single-segment firms, which is consistent with the general notion that the economy is becoming more specialized: While in 1990 about 47% of book assets in the U.S. economy were held by single-segment corporations, this number jumps to 63% in 2011.

Our paper develops a real-options model of diversification in the spirit of Hackbarth and Morellec (2008), where conglomerates can reallocate technologies/resources optimally across divisions, thus generating synergies. The key feature of the model is that synergies depend on the level of technological specialization, which therefore determines the patterns of corporate-diversification activity. In particular, our model has two main implications. First, optimal technological diversity across divisions decreases with technological specialization, leading to more-focused conglomerates in equilibrium. Second, the benefits of ex-post resource reallocation decrease as the economy becomes technologically more specialized, which leads to a gradual reduction in corporate diversification. A calibrated version of the model

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

²Smith (1776).

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

matches the trends in conglomerate focus and pervasiveness, as well as several other empirical magnitudes: growing output, level and trend of the diversification discount, frequency and returns of diversifying mergers, and frequency of refocusing activity.

We model an economy that is populated by *business units*, which are taken to be the elementary agent of production. Time is continuous, and single-segment firms can engage in diversifying mergers.⁴ Following Rhodes-Kropf and Robinson (2008), mergers are modeled in the spirit of search-and-matching literature on unemployment (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.⁵ Once a conglomerate becomes inefficient, it refocuses with some probability, also according to an exogenous Poisson process.

In our model we employ a broad concept of “technology”, which includes not only technical capabilities, but also a firm’s managerial/organizational know-how. Furthermore, we model production technology and diversification synergies using a spatial representation. Specifically, each business unit is characterized by a location on a technology circle. Business units pursue projects, which are also characterized by a location on the circle, representing the ideal business unit (or technology type) to undertake the project. Business units randomly draw projects within a neighborhood of their technology, and output is decreasing in project-business-unit distance. Business units thus face the risk of drawing a project for which they are ill-equipped, which motivates corporate diversification. Diversifying mergers generate synergies because business units within the same firm are allowed to trade projects whenever this is efficient; this in-house project trade represents within-conglomerate re-

⁴For 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).

⁵This 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).

source reallocation. Thus our approach is close to the internal capital markets literature (Stein, 1997; Scharfstein and Stein, 2000), albeit we consider an ability to reallocate technological capabilities rather than financial capital. An implicit assumption of our model is that such reallocation is feasible within firms but not across firms, for example because of greater adverse selection.⁶

In our spatial model, technological specialization refers to the range of project types business units face. In periods of low specialization this range is wide, which implies corporate diversification can add much value through ex-post reallocation. As specialization increases, business units experience a higher frequency of projects for which they have a comparative advantage, with two implications: average output increases and diversification synergies become lower.

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 segment distance, driven by the following trade-off. On one hand, diversifying synergies initially increase in segment distance, or technological diversity. The intuition for this effect is that complementarity is relatively low if two business units are very similar, since trading projects in that case can only generate limited gains (in fact zero as technologies fully overlap). 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.

Using data on corporate-diversification activity in the U.S., we then perform a calibration of our dynamic model. In data, we measure the distance across conglomerate segments using as a topology an inter-industry network based on input-output flows. The calibration employs a growth rate for technological specialization that generates reasonable output growth, and we are able to match important magnitudes that characterize aggregate corporate-

⁶This assumption is in line with an interpretation of the boundaries of the firm as information boundaries, as suggested, for example, in Chou (2007).

diversification activity: the proportion of assets allocated to single-segment firms in the economy, average announcement returns of diversifying mergers, and the so-called “diversification discount”. We note that although we match the diversification discount, this discount is only apparent, since firms are perfectly aligned with shareholder-value maximization at the time that mergers take place.⁷

Our calibrated model explains not only levels, but also corporate-diversification trends, although we only partially match the average growth rate in segment distance (the model-implied magnitude is at most three-quarters of the absolute growth rate in data). The calibration also matches two other trends in data, namely an increase in the Tobin’s Q of single-segment firms and an increase in conglomerate excess value, an industry-adjusted valuation measure. The calibration matches the aforementioned empirical patterns while using a standard level for the discount rate, reasonable frequencies of merger and refocusing activity for the representative firm, and a reasonable average level for Tobin’s Q .

We also investigate the model’s cross-sectional implications. 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.

The empirical finding that excess value increases with segment distance stands in contrast with the mainstream stance in finance research about *relatedness* (broadly defined), which is usually understood to be a positive factor behind synergies (Berger and Ofek, 1995; Fan and Lang, 2000; Hoberg and Phillips, 2010; Bena and Li, 2013). However, a positive association between relatedness and value is potentially identified by unrelated deals that are

⁷Our explanation for the diversification discount is in the spirit of Anjos (2010). Other papers have proposed rational explanations for the discount using dynamic models; see for example Matsusaka (2001), Bernardo and Chowdhry (2002), Maksimovic and Phillips (2002), and Gomes and Livdan (2004).

motivated, for example, by managerial empire-building; and not all empirical measures of similarity/relatedness necessarily pick up such agency effects to the same extent. Therefore, these two views are not necessarily inconsistent or mutually exclusive.

In summary, our paper provides the following contributions to the finance literature. First, we provide a novel, network-based empirical measure of technological diversity, which uses the overall inter-industry architecture of the economy. Second, we document novel empirical facts about the evolution of corporate-diversification activity. Third, we develop a novel theory explicitly linking technological specialization and the diversification synergies that accrue from within-firm resource reallocation. Fourth, the calibrated version of our model quantitatively matches the empirical patterns of corporate diversification.

The remainder of the paper is organized as follows. Section 2 presents some motivating evidence on the evolution of conglomerate activity. Section 3 develops the theoretical setup, which entails a model for the relationship between technological specialization, segment distance, and flow synergies from corporate diversification; and a model for the process through which diversification activity occurs and firm boundaries change. Section 4 performs a calibration exercise. Section 5 investigates the model's cross-sectional implications. Section 6 concludes. An appendix contains all proofs, an extension to the main model, summary statistics, and details on variable construction and model implementation.

2 Motivating evidence

This section presents some initial evidence on the evolution of corporate-diversification activity. Detailed summary statistics are presented in the appendix (section A.5).

2.1 The evolution of segment distance

The level of *relatedness* across segments has been a key variable in the study of conglomerates (Berger and Ofek, 1995; Fan and Lang, 2000; Custódio, 2013). One of the contributions of

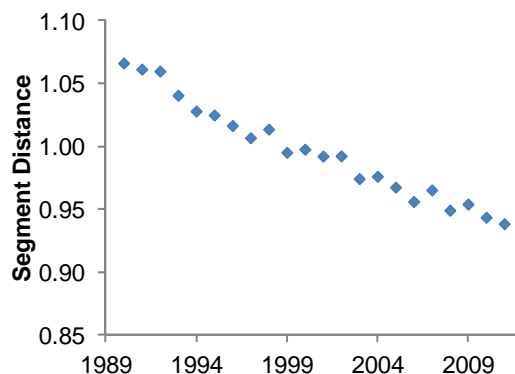


Figure 1: Segment Distance over Time. The figure shows average segment distance, for the period 1990-2011. Segment Distance is the average input-output-based distance across conglomerate segments. Details on the construction of segment distance are presented in the appendix (section A.1).

our paper is a novel measure of (un)relatedness, which we term segment distance, that captures the level of technological diversity across conglomerate divisions. 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;⁸ and finally, for a particular conglomerate, we identify all relevant industry pairs and compute their average distance.⁹ In the appendix we provide details about the construction of the segment-distance variable (section A.1).

The empirical evolution of segment distance is quite uncontroversial and intuitive: Figure 1 shows that for the period 1990-2011 there was a gradual, almost linear decrease in segment distance.¹⁰ The trend is the same irrespective of whether we look at averages or medians: Segment distance for a representative conglomerate dropped about 12% over a 21-year period. The slow gradual decline in segment distance is consistent with a view that technological specialization is slowly but steadily increasing in the economy, and our model provides a rigorous formalization for this intuition.

We view segment distance as a proxy for the level of technological diversity across con-

⁸Our approach to converting the U.S. input-output matrix into a network follows Anjos and Fracassi (2014) closely.

⁹Fan and Lang (2000) also propose relatedness measures based on input-output flows, but do not consider the overall network architecture, which we do.

¹⁰We start our data in 1990 because we require NAICS classification codes in order to construct the input-output-based industry network.

glomerate divisions. There are three main advantages to segment distance, compared to other relatedness measures: First, it is defined for all industries in the economy, and not just the subset of manufacturing industries.¹¹ Second, our concept of “technology” is quite broad, as in standard macroeconomic models, and includes a firm’s managerial/organizational technology, which is potentially similar for industries that are close-by in the economy-wide supply chain.¹² Finally, our segment-distance variable also has the advantage of not being overly dependent on the specific industry-classification scheme, unlike the one proposed by Berger and Ofek (1995). In particular, if two industries are focusing on a similar economic activity, one would expect, everything else constant, that these two industries have a similar set of customer and supplier industries. Sharing these indirect connections yields a low segment distance, which thus is capturing how equivalent two industries are in the economy-wide supply chain. Moreover, segment distance generalizes this notion of technological equivalence by also including higher-order indirect connections—customers of customers, customers of suppliers, and so on.

2.2 Additional trends

This section documents additional time-series patterns that will also be accounted for by our model.

First we turn to the pervasiveness of corporate-diversification activity. The top panel of figure 2 shows the evolution of the proportion of book assets allocated to single-segment companies. We find a clear positive trend, even though the data is noisy and apparently cyclical. This is partly due to underlying economic forces, but also a consequence of the change in segment-reporting requirements introduced in 1997-1998.¹³ The bottom-left panel

¹¹This is important for our purpose of characterizing economy-wide corporate-diversification activity, and so we would not want to employ a technological similarity measure that is only defined for manufacturing, as for example in Bena and Li (2013).

¹²For example, suppose two vertically-disconnected industries A and B share a key supplier industry C; then it seems reasonable that a management team of company A would be relatively efficient in managing firm B.

¹³From SFAS 14 to SFAS 131 (see Sanzhar, 2006 for more details about the rule changes).

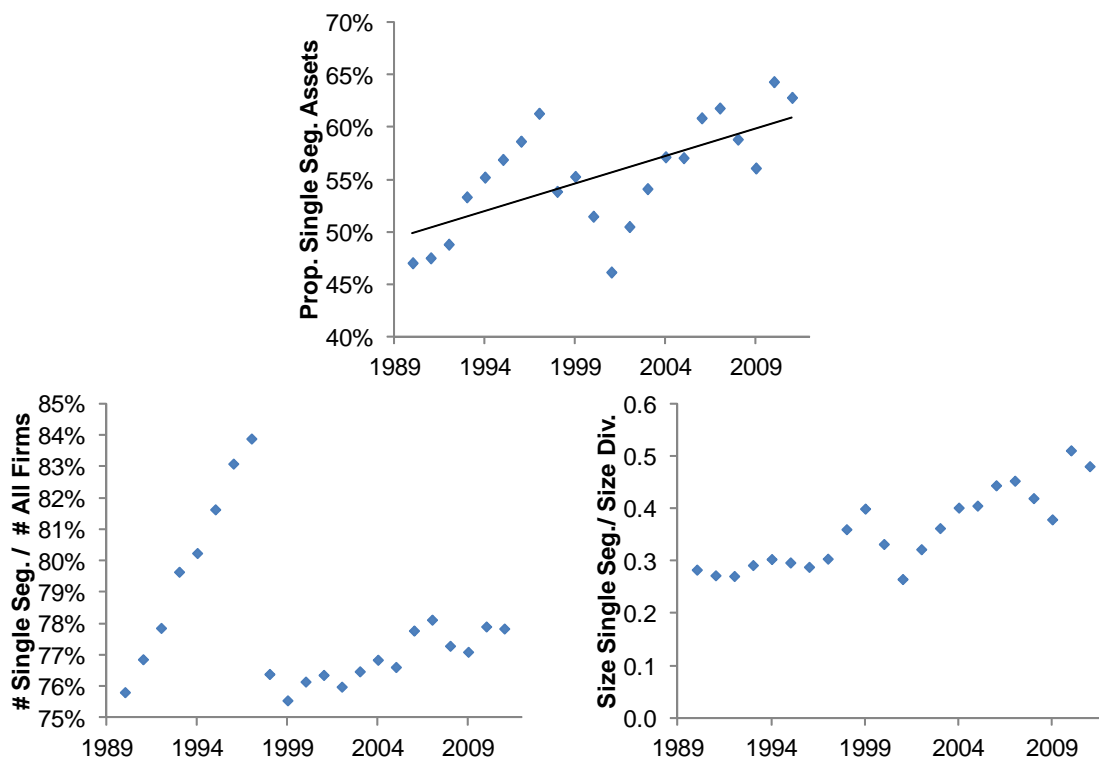


Figure 2: Pervasiveness of Single-Segment Firms. The top panel shows the proportion of total assets in the economy allocated to single-segment firms (and a linear trend line). The bottom-left panel shows the fraction of firms that are single-segment, for the period 1990-2011. The bottom-right panel shows the size ratio between single-segment and diversified firms.

of figure 2 plots the fraction of firms classified as single-segment. There is a clear discontinuity in 1998, consistent with the change in reporting requirements. For each subperiod, the left panel shows a clear positive trend, albeit the trend is suspiciously strong for early years.¹⁴ The bottom-right panel of figure 2 plots the average asset-size ratio of single-segment to diversified corporations, where a clear upward trend is present. In summary, we believe this evidence indicates a generalized increase of single-segment activity in the economy, which is also consistent with the notion of ever-increasing technological specialization.

We conclude our characterization of corporate diversification by analyzing valuation trends for both single-segment and diversified firms. The left panel of figure 3 shows the evolution of Tobin's Q for single-segment firms, with a clear positive trend. We also note that other authors have suggested a long-term increase in Tobin's Q (see Obreja and Telmer,

¹⁴This may be related to an attempt by some conglomerates to try to appear as single-segments, in line with Sanzhar (2006).

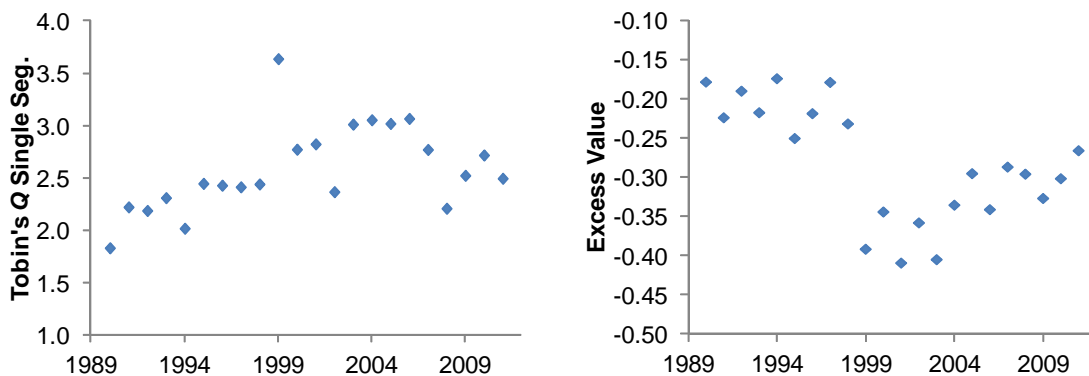


Figure 3: Evolution of Valuation Measures. The left panel shows the average Tobin's Q of single-segment firms for the period 1990-2011. The right panel shows conglomerate excess value, which is defined as the log-difference between the Tobin's Q of a conglomerate and the Tobin's Q of a similar portfolio of single-segment firms, following Berger and Ofek (1995).

2013).

The right panel of figure 3 plots excess value, that is, the log-difference between the value of the conglomerate and the value of a comparable portfolio of single-segment firms.¹⁵ As in other papers on corporate diversification, average excess value is negative (the celebrated diversification discount). Excess value for the representative conglomerate exhibits a strong discontinuity around the introduction of the new segment-reporting requirements. In the first sub-period there is no apparent trend in excess value, which could potentially be explained by the fact that many single-segment firms were actually misclassified conglomerates. Inclusion of conglomerates in the single-segment sample could make the excess-value variable very noisy (and potentially biased), obscuring any eventual trend. The second subperiod shows a clear upward trend in excess value.

3 Model

In the previous section we documented several empirical patterns. In particular, there is strong evidence that, over time and for the period 1990-2011, (i) diversified firms tend to exhibit lower segment distance; and (ii) the proportion of assets allocated to single-segment firms is increasing. The evidence also suggests that single-segment Q increases, but

¹⁵Excess value was originally introduced by Berger and Ofek (1995) and is extensively used in the diversification literature.

conglomerate Q increases even more.

We now turn to developing our theoretical framework, which will offer an explanation for the observed trends. We start by developing a static equilibrium model for flow payoffs (section 3.1), which we then embed in a dynamic search-and-matching framework (section 3.2).

3.1 Flow payoffs

The economy comprises a continuum of business units (henceforth BUs), where BU i is characterized by a location α_i on a circle with measure 1, represented in figure 4.¹⁶ 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 3.2. The next two subsections further characterize the flow payoffs of single-segment and diversified firms.

3.1.1 Single-segment firms

Each BU in the economy undertakes one project,¹⁷ 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 4. The higher σ is, the higher the risk that business units are presented with

¹⁶The 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.

¹⁷An implicit assumption of our model is that projects cannot be traded across firms. This could be due, for example, to adverse selection; and would be consistent with interpreting the boundaries of the firm as information boundaries (as suggested, e.g., in Chou, 2007).

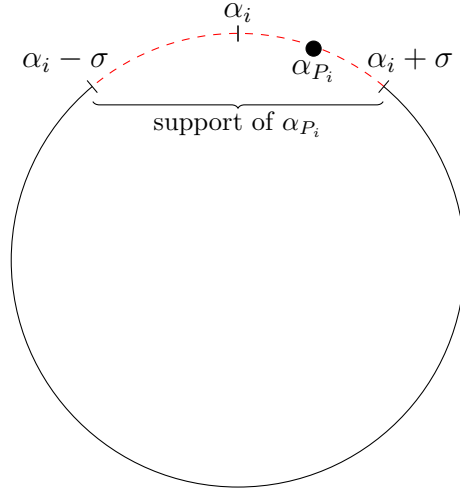


Figure 4: Technologies and Projects: Spatial Representation. The figure depicts a circle where both projects and business units are located. The location of the business unit (α_i) represents its technology, whereas the location of projects (α_{P_i}) represents the ideal technology to undertake that particular project. The figure also shows that business units draw projects from locations close to their technology.

projects for which they are ill-equipped, and we interpret the inverse of σ as the degree of *technological specialization*. Specialization in our model thus refers to the extent to which business units are able to find good projects for their technology, which is consistent with the fundamental notion that an increase in focus delivers higher productivity. In particular, we assume that σ gradually decreases over time, which translates into positive economic growth (dynamics are detailed in section 4.2). For tractability we assume $\sigma < 1/4$, which greatly simplifies the analysis.¹⁸

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

$$\pi_i = 1 - \phi z_{i,P_i}, \quad (1)$$

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

¹⁸Tractability with low enough uncertainty about project location originates from the fact that we only have to consider one-sided overlap in project-generating regions. The advantage of this assumption is clear in the derivations and proofs presented in the appendix. We also believe this assumption is fairly innocuous in terms of the main results.

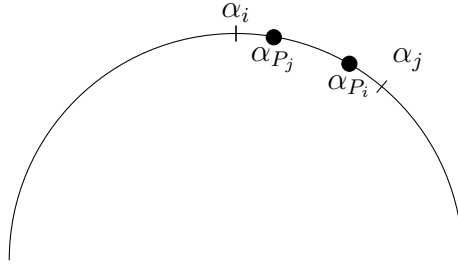


Figure 5: Conglomerates and Reallocation: Spatial Representation. The figure depicts the location of conglomerate segments on the technology circle; and shows an instance where projects are optimally swapped across segments, i.e. division i is assigned to project j and vice-versa.

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

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

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” ($\sigma = 0$).

3.1.2 Diversified firms

To keep the framework tractable, the only form of corporate diversification we consider is a conglomerate with two segments. If BU i is part of the same firm as BU j , then profits are similar to those of a single-segment firm, with the exception that projects can be traded (swapped) inside the firm; and this ex-post choice is assumed to be made optimally by the headquarters of the multi-segment firm so as to minimize the total costs of project-technology misfit (represented in figure 5). 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.

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. For now we take the proportions of good and bad conglomerates as given; these are endogenized later (section 3.2). 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.

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:*

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

Figure 6 depicts the relationship between segment distance and average division profits, and illustrates the natural ambiguity in this relationship. If distance is too low, there are many efficient project transfers, however the average gain of each transfer is small. If distance is too high, then realized project transfers correspond on average to a large gain; however, each division is usually the closest to the projects it generates, and so transfers are rare. The 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 , which is intuitive.

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

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

with associated expected BU profit of

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

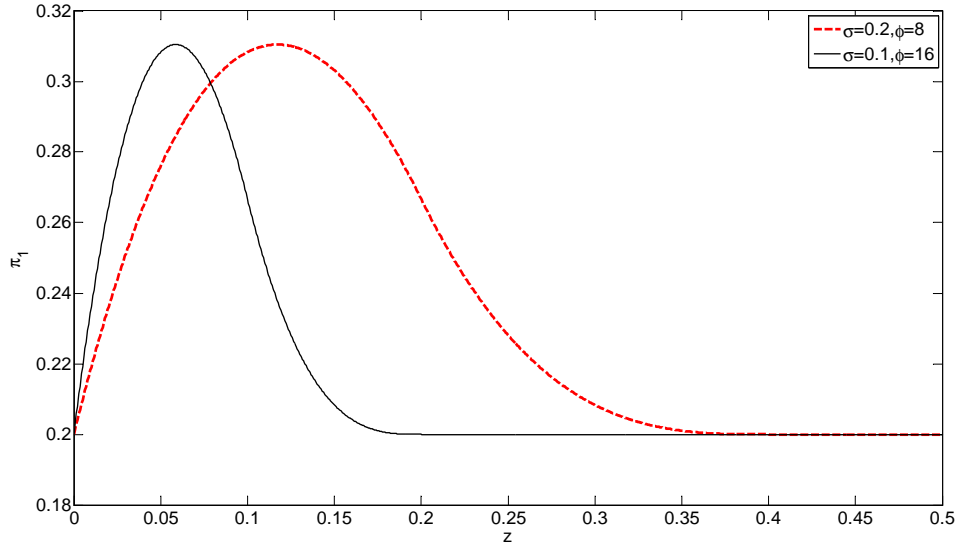


Figure 6: Segment Distance and (Static) Profits. The figure plots profits π_1 as a function of segment distance z .

If σ is interpreted as a measure of the inverse of specialization, then an increase in specialization (lower σ) would imply that diversified firms should become more specialized too, that is, one should observe most conglomerates with segments that are closer or less diverse. This would be consistent with the empirical pattern we documented in section 2 (figure 1).

Inspecting figure 6, it is ambiguous which empirical 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 would 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).

Comparing the two plots in figure 6 one observes that the relationship between segment distance and profits is *scaled* by σ . As long as the product $\phi\sigma$ is constant, the maximal value

of synergies is the same (see proposition 2). Therefore, holding the product $\phi\sigma$ constant, it would not be possible to distinguish between an economy where σ is high and the distribution of firms has wide support (dashed curve of figure 6) from an economy with low σ but where the distribution of firms has narrow support (solid curve of figure 6). This point is important for our calibration, where given the argument just outlined we set the initial σ at an arbitrary level.

Bad conglomerates

As will become apparent later, matching data requires the existence of some additional costs associated with corporate diversification. 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.

3.2 Dynamics

3.2.1 Matching technology

We now complete our setup, by considering a dynamic continuous-time economy comprising a continuum of infinitely-lived business units (BUs) uniformly located on the circle of 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.¹⁹ 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 , and we choose parameters such that it is efficient to break a bad conglomerate apart. Bad conglomerates refocus according to a Poisson process with intensity λ_2 .

An important ingredient of the model is how to specify the segment distance at which matches occur. With the caveat that equilibrium has not yet been defined, as long as one focuses on symmetric equilibria then it makes sense that the matching technology be independent of specific locations in the circle. Based on this rationale, 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 $[0, 1/2]$.

3.2.2 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

¹⁹Our approach is similar to Hackbarth and Morellec (2008), who develop and calibrate a real-options model of mergers.

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 does not lend itself to being represented with simple equations. The steady-state case thus provides a useful benchmark to understand the basic mechanics of the model. In the appendix we detail the solution to the more-general case (section A.3).

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:

$$J_t = \sup_{\{\tau\}} \left\{ \mathbb{E}_t \left[\int_{u \in [t, +\infty] \cap \{\tau, \tau_2\}} e^{-r(u-t)} [\pi_1(z_{\sup\{\tau < u\}}) - \beta \mathbb{1}_{\sup\{\tau < u\} < \sup\{\tau_1 < u\}}] du + \int_{u \in [t, +\infty] \setminus \{\tau, \tau_2\}} e^{-r(u-t)} \pi_0 du \right] \right\}, \quad (6)$$

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 distance of the two divisions inside the diversified firm.

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

Definition 1 (*Equilibrium*) *A Markov Perfect Equilibrium of this economy is characterized by an unchanging proportion of single-segment firms $p \in [0, 1]$, a 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 it is the case that the merger acceptance policy solves optimization problem (6).*

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

Proposition 3 *In an equilibrium with no mergers, the value of single-segment firms J_0 is equal to π_0/r . In an equilibrium with mergers, the optimal policy of single-segment firms is characterized by accepting matches with segment distance in an interval $[z_L, z_H]$. In such an equilibrium, the time- t value of a business unit inside a bad conglomerate, J_2 , is a simple function of the segment distance at which the merger took place (z):*

$$J_2(z) = \frac{\pi_1(z) - \beta + \lambda_2 J_0}{r + \lambda_2} \quad (7)$$

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

$$J_1(z) = \frac{\pi_1(z)(r + \lambda_1 + \lambda_2) - \lambda_1 \beta + \lambda_1 \lambda_2 J_0}{(r + \lambda_1)(r + \lambda_2)}. \quad (8)$$

The value of single-segment firms J_0 is characterized as

$$J_0 = \frac{\pi_0(r + \lambda_1)(r + \lambda_2) + \lambda_0 q(r + \lambda_1 + \lambda_2)\bar{\pi}_1 - \lambda_0 q \lambda_1 \beta}{(r + \lambda_0 q)(r + \lambda_1)(r + \lambda_2) - \lambda_0 q \lambda_1 \lambda_2}, \quad (9)$$

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

$$q := \frac{z_H - z_L}{0.5} \quad (10)$$

$$\bar{\pi}_1 := \int_{z_L}^{z_H} \frac{1}{z_H - z_L} \pi_1(z) dz \quad (11)$$

Equation (9) describes the equilibrium value of single-segment firms, which embeds the value of the option to diversify. It is also clear in equations (7)-(9) how 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

$$p = \frac{1}{1 + \lambda_0 q (1/\lambda_1 + 1/\lambda_2)}. \quad (12)$$

2. The fraction of bad conglomerates is

$$w = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \quad (13)$$

3. There exists a threshold C , defined as

$$C := \frac{6\lambda_1\beta}{(\sqrt{2} - 1)(r + \lambda_1 + \lambda_2)}, \quad (14)$$

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 or λ_2 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 obtains because 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.²⁰

The third result in proposition 4 shows that mergers only take place if either the location of projects is highly uncertain (high σ) or the cost of project-technology misfit is high (ϕ), 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, gaged 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; hence the appearance of this parameter on the RHS of (13). Finally, when $\lambda_1 \rightarrow 0$, organizational-complexity costs no longer factor into the diversification trade-off (RHS of (13) becomes zero), 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.*

4 Calibration

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 .

4.1 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

²⁰This argument is along the lines of Anjos (2010).

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	0.37
Likelihood of becoming bad conglomerate	λ_1	0.09
Likelihood of refocusing	λ_2	0.16
Overhead cost of bad conglomerates	β	0.40
Cost of project technological mismatch	ϕ	8.50
Inverse of technological specialization	σ	0.20

the steady-state model is inadequate to provide implications about how changes in specialization (σ) affect corporate-diversification *trends*,²¹ 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 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), 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 , λ_2 , and σ . We set the discount rate r at 10%, which seems reasonable for the average firm in the economy. As for λ_2 , we set it so as to obtain a reasonable rate of refocusing. In our data, the fraction of conglomerates reducing the number of segments over a one-year period is 15%; to match this frequency of refocusing we therefore need

$$1 - e^{-\lambda_2} = 0.85,$$

which implies $\lambda_2 = 0.16$. Finally, we set $\sigma = 0.2$, which is just a normalization. As explained in section 3.1, it would not be possible in our model to separately identify σ from the ϕ .²²

We use five moments in data as targets for calibrating the remaining four parameters.

²¹The only alternative would be a comparative-statics exercise, which would not factor in the fact that firms presumably know that σ is changing.

²²See figure 6 and related text.

We describe the rationale for each choice below:

- In data, the average Tobin's Q of single-segment firms is 2.6. We want to obtain J_0 that is close to this but we note that there is no cash flow growth in our steady-state model, so it seems natural to target a relatively more conservative magnitude. If we added constant growth to our model, say at 2% per annum, then a Tobin's Q of 2 with no growth is comparable to

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

which is close to 2.6.

- Our data counterpart to p , the fraction of single-segment firms in the economy, is the in-sample average proportion of book assets owned by single-segment corporations, approximately 55%.
- We match the model-implied excess value to its counterpart in data, which in our sample is -0.28 . In the model, excess value is easily computed from equations (7)-(9) and (13):

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

- 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). In the model, this likelihood corresponds to 1 minus the probability that the firm does not engage in any merger, which is given by

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

- Finally we attempt to match the average magnitude of diversifying-merger announce-

Table 2: Model outputs and data (1/2). The table shows key moments, both in the calibration and in data; for the steady-state calibration. “Single-Seg. Value” is the Tobin’s Q of single-segment firms; “Prop. Single-Seg.” is the proportion of assets in the economy allocated to single-segment firms; “Av. Excess Value” is the unconditional average excess value of conglomerates; “Probab. of M&A” stands for the likelihood that a single-segment BU engaged in at least one merger deal; and “Av. Div. Returns” stands for the average announcement returns of diversifying mergers.

Moment	Model Counterpart	Calibration Output	Data/target
Single-Seg. Value	J_0	1.53	2.00
Prop. Single-Seg.	p	50%	55%
Av. Excess Value	$\frac{wE[J_2]+(1-w)E[J_1]-J_0}{J_0}$	-0.24	-0.28
Probab. of M&A	$1 - e^{-\lambda_0 q}$	5.6%	6.0%
Av. Div. Returns	$\frac{E[J_1]-J_0}{J_0}$	3.5%	3.8%

ment returns, which in the model is simply

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

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).

Table 1 summarizes the choice of parameters, and table 2 reports key moments. The procedure we use for generating parameters is to minimize the equally-weighted sum of squared (relative) differences between model and data.²³ The calibration yields a reasonable fit to data, in particular in terms of two key corporate-diversification magnitudes: how many conglomerates there are and how discounted they appear to be relative to single-segment firms.

4.2 Time-varying technological specialization

This section builds on the steady-state calibration, adding a time-varying σ . Our final objective is to compare model outputs with the corporate-diversification data presented in

²³For each moment, the penalty function is thus $[(target - output)/target]^2$.

section 2: decreasing segment distance (figure 1), growing fraction of single-segment firms (figure 2), increasing single-segment Tobin's Q (figure 3), and increasing excess value (figure 3).

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 -0.3% , 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).

Now we turn to the dynamic implications of our calibration. The key outputs are illustrated in figure 7 for the period 1990-2011; outputs for a longer period of time are presented and discussed in the appendix (see figure A.2).

The top-left panel of figure 7 shows that a decrease in σ , which we interpret as an increase in specialization, leads to a higher proportion of single-segment firms. This is in line with the trend in data, and the intuition for the result is straightforward: as σ reduces, the benefits of combining non-redundant technologies are lower relative to the potential costs of organizational complexity, and thus in equilibrium one observes fewer conglomerates. The top-right panel shows how a decrease in σ over time leads to a decrease in segment distance for the average conglomerate, also in line with data. The result follows from the fact that a lower σ implies a narrower optimal range for M&A activity, as explained in section 3.1.2. The bottom-left panel shows that the value of single-segment firms increases as σ is reduced, which follows directly from the fact that σ gages the average level of project-firm misfit. Finally, the bottom-right panel of figure 7 shows that excess value increases for higher levels of specialization. To explain this result, we start by noting that as σ decreases, both the

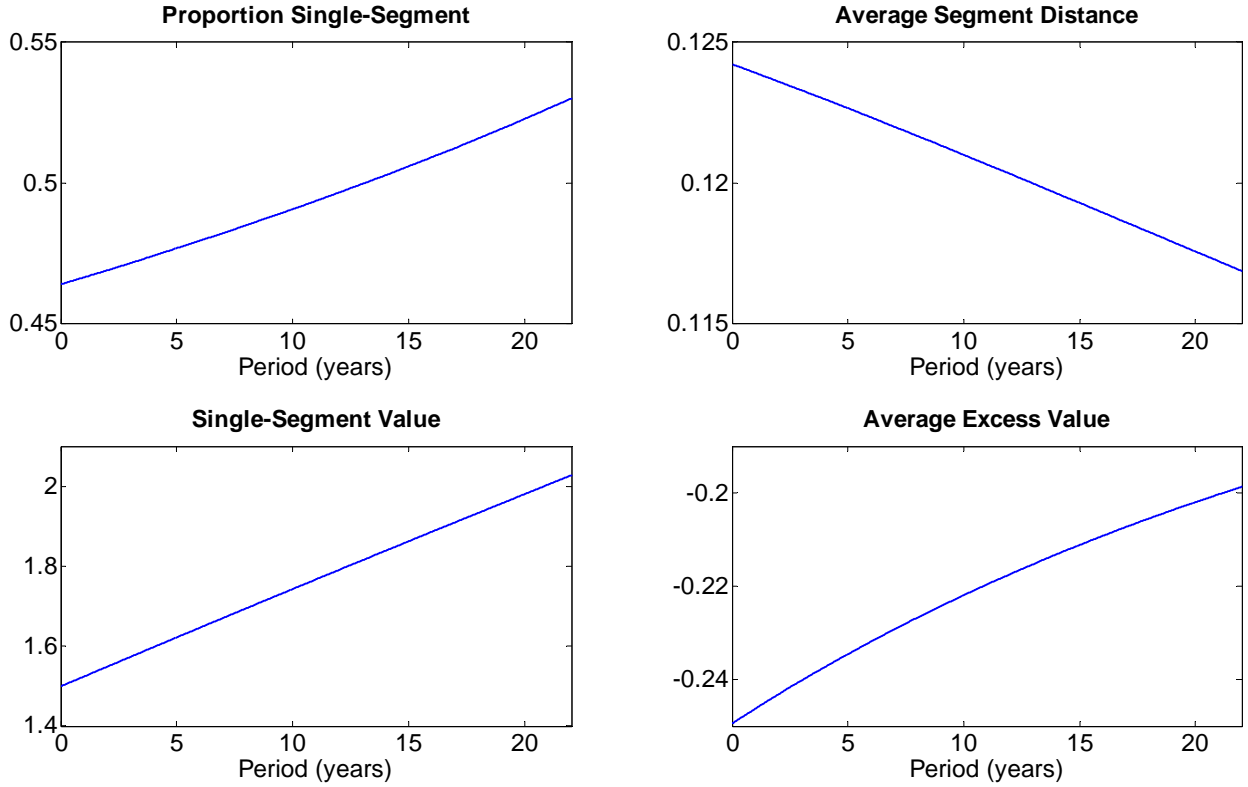


Figure 7: Calibration with Time-Varying Specialization: Key Outputs. The top-left panel shows the proportion of single-segment assets in the economy; the top-right panel shows the average diversified-firm segment distance; the bottom-left panel plots the value of single-segment firms; and the bottom-right panel plots conglomerate excess value.

value of single-segment firms and diversified firms increases. This effect is independent of organizational-complexity costs (β), and so in relative terms the value of bad conglomerates increases by a significant percent amount. If there are enough bad conglomerates in the economy, and/or if the costs of organizational complexity are high, then a decrease in σ is thus followed by an increase in excess value. This mechanism implies that we would not observe an increase in excess value if there was no diversification discount, since in such a setting percent increases in conglomerate value would be low.²⁴

So far we have shown that the dynamic predictions of the model are in line with data, at least qualitatively. Next we turn to a more quantitative assessment, and below we elaborate on the rationale for each data target:

²⁴Indeed, if we choose parameters such that there is no diversification discount (low β and/or high λ_2), then average excess value actually decreases over time. For the sake of space these results are not shown.

Table 3: Model outputs and data (2/2). The table compares the annual average growth rates implied by the model for each variable, and compares it to a target interval in data. J_0 is the value of single-segment firms, $|EV|$ is absolute average excess value, p is the fraction of single-segment firms, and \bar{z} is average segment distance.

Variable	Model-implied growth rate	Data target
p	0.6%	[0%, 3%]
\bar{z}	-0.3%	[-0.9%, -0.4%]
J_0	1.4%	[1%, 5%]
$ EV $	-1.0%	[-5%, 2%]

- **Fraction of assets within single-segment firms.** If we consider all data points from the top panel in figure 2, the growth rate for this variable has an in-sample mean of 1.6% p.a., with a standard error of about 1.5%. If we focus on the period after 1998, which given the classification issues raised by Sanzhar (2006) seems reasonable, then the average growth rate is about 0.5% p.a., with a standard error of 2.1%. In light of these computations, we believe an interval of [0%, 3%] is appropriate as a target for the model.
- **Segment distance.** Inspection of figure 1 shows that this time series is relatively smooth. The average growth rate in segment distance is -0.6% (-0.68%) p.a. if we take the average (median), with a standard error of about 0.18% (0.20%). Based on these magnitudes, we define a reasonable target interval for the growth rate of segment distance as [-0.9%, -0.4%].
- **Value of single-segment firms.** The data for the Tobin's Q of single-segment firms, shown in the left panel of figure 3, is quite noisy. Focusing on the entire period, the average growth rate for this variable is about 2.7% (1.6%) p.a. if we take the average (median), with a standard error of about 3.7% (2.9%). Based on these magnitudes, and also the fact that other authors suggest Tobin's Q has been increasing over time,²⁵ we define a reasonable target interval for the growth rate of single-segment Tobin's Q as [1%, 5%].

²⁵See Obreja and Telmer (2013).

- **Excess value.** For this magnitude, and due to the classification concerns raised by Sanzhar (2006), we focus on the more-recent observations (post-1998). The average growth rate for *absolute excess value* over this period is about -2.3% (-0.9%) p.a. if we take the average (median), with a standard error of about 3.9% (3.2%). Based on these magnitudes, we define a reasonable target interval for the growth rate of absolute excess value as $[-5\%, 2\%]$.

Table 3 compares model outputs and data. The model fares relatively well in all dimensions, albeit there is a slight mismatch in terms of the growth rate of segment distance: the model-implied magnitude of -0.3% is larger than the upper bound for the data target (-0.4%).

5 Cross-sectional implications

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 3.1 (see figure 6), 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.

5.1 Reduced-form evidence

The left panel of figure 8 describes the segment-distance distribution for our whole sample, covering the period 1990-2011. Consistent with the prediction of our theory, we observe conglomerates cluster at intermediate distances. The right panel of figure 8 shows the empirical association between segment distance and conglomerate valuation. Here we should also observe a non-monotonic relationship, but the relationship is linear and positive. In section 5.2 we address this mismatch between theory and data. We also find that the positive as-

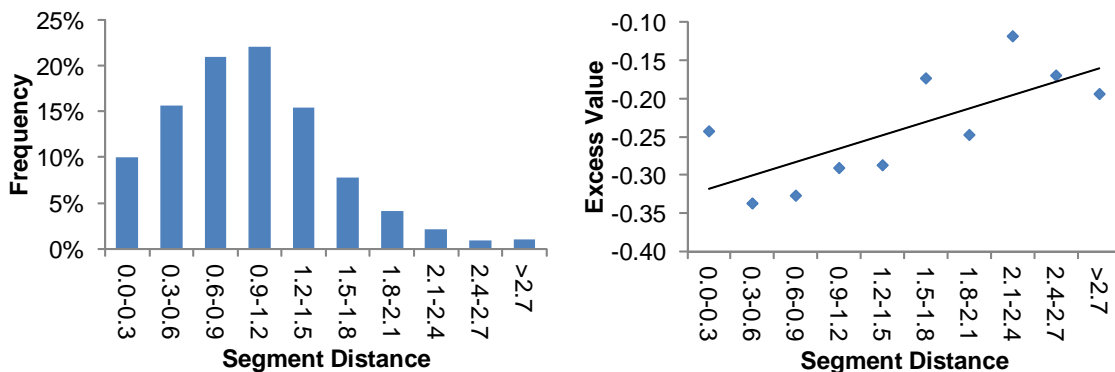


Figure 8: Segment Distance and the Cross Section of Conglomerates. The left panel shows the segment-distance distribution. The right panel shows conglomerate average excess value, conditional on segment-distance class. Excess value is defined as the log-difference between the Tobin’s Q of a conglomerate and the Tobin’s Q of a similar portfolio of single-segment firms, following Berger and Ofek (1995). Segment Distance is the average input-output-based distance across conglomerate segments.

sociation between segment distance and excess value is robust to controlling for many other factors, as shown in table 4. For ease of interpretation, all variables have been standardized.

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 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 (2014), 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),²⁶ and specification (4) includes firm fixed effects, which allows us to rule out an explanation based on persistent managerial skill or unobserved organizational capital, where

²⁶Results are however similar if we use raw financial conglomerate variables, instead of computing excess measures.

Table 4: Excess Value and Segment Distance. The dependent variable is Excess Value, defined as the log-difference between the Tobin's Q of a conglomerate and the Tobin's Q of a similar portfolio of single-segment firms, following Berger and Ofek (1995). The table presents ordinary least squares regression coefficients and robust t-statistics clustered at the conglomerate level. The main explanatory variable is Segment Distance, defined as the average level of binary distance for every possible pair of industries that the conglomerate participates in, using the 6-digit Input-Output industry classification system. All variables are defined in detail in the appendix. A constant is included in each specification but not reported in the table. 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	0.043*** (2.80)	0.037** (2.07)	0.035* (1.95)	0.084*** (3.53)
N. Segments		-0.054*** (-3.10)	-0.063*** (-3.42)	-0.080*** (-3.90)
Related Segments		0.051*** (2.74)	0.043** (2.38)	0.016 (0.82)
Vert. Relatedness		0.023* (1.90)	0.017 (1.34)	0.073*** (3.05)
Excess Centrality		0.040** (2.13)	0.036* (1.90)	0.062** (2.27)
Excess Assets			0.055*** (2.78)	-0.059 (-1.54)
Excess EBIT/Sales			-0.091*** (-9.18)	-0.026*** (-2.70)
Excess Capex/Sales			0.015*** (3.02)	0.029*** (8.73)
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	No	Yes
R^2	0.015	0.018	0.030	0.028
Obs.	22,425	22,425	21,516	21,516

better firms are the ones that simultaneously are more profitable running their businesses and also have more ability to evaluate merger/expansion opportunities at a distance.²⁷

Segment distance has an economically-significant impact in terms of conglomerate value. A one-standard-deviation increase in segment distance is associated with an increase of between 0.035 and 0.084 standard deviations in excess value. Excess value has a standard deviation of 0.66, so this corresponds to an increase of between 0.023 and 0.055 in excess

²⁷With the caveat that time-varying managerial skills or firm organizational capital could still render our results spurious.

value, that is, between $0.023/0.72 \approx 3.2\%$ and $0.023/0.72 \approx 7.6\%$ of firm value for the average conglomerate.

In table 4 the coefficients on number of segments, related segments, and vertical relatedness are all consistent with previous literature: relatedness is associated with higher firm value. This begs the question of why the results are qualitatively different with segment distance and excess centrality. Our theory notwithstanding, it is certainly plausible that 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 non-redundant information (excess centrality).

5.2 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 left panel of figure 8, is consistent with the notion that firms prefer intermediate-distance combinations to low-distance combinations. That the frequency afterwards decreases is however not necessarily a function of firms not preferring high-distance deals, *per se*. In particular, it seems reasonable that fewer M&A deals are free from serious adverse-selection issues as distance increases (explaining the low frequency); but, for those where adverse selection is

indeed not a concern, then one observes relatively high synergies (explaining high Tobin's Q for high-segment-distance firms). 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.

6 Conclusion

Our paper contributes to the literature on corporate diversification in several ways. First we develop a novel theory of conglomerates, explicitly linking the seminal concept of technological specialization to corporate-diversification activity. Specifically, we show how it is optimal for the divisions within a conglomerate to be technologically more similar as technological specialization increases, and also how technological specialization leads to the existence of fewer conglomerates. Second, we provide novel empirical facts about the evolution of corporate diversification in the U.S., and show that the key predictions of the model are borne out in data: there is a salient, steady trend towards conglomerates that are more focused/related; and the fraction of assets owned by diversified firms is decreasing over time. Our calibrated model also matches data in other dimensions, namely in terms of the level and trend of the diversification discount, the frequency of diversifying mergers and refocusing activity, and the aggregate Tobin's Q . Finally, our paper develops a novel empirical approach to measuring relatedness across conglomerate segments, which builds on the economy-wide inter-industry trade network.

References

- 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, 2014, Shopping for information? Diversification and the network of industries, Working paper, available at SSRN.
- Bena, Jan, and Kai Li, 2013, Corporate innovations and mergers and acquisitions, *Journal of Finance* (forthcoming).
- Berger, Philip G., and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–65.
- Bernardo, Antonio E., and Bhagwan Chowdhry, 2002, Resources, real options, and corporate strategy, *Journal of Financial Economics* 63, 211–234.
- Burt, Ronald S., 2005, *An Introduction to Social Capital* (Oxford University Press).
- Campa, José Manuel, and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 1731–1762.
- Chou, Eric S., 2007, The boundaries of firms as information barriers, *RAND Journal of Economics* 38, 733–746.

- Custódio, Cláudia, 2013, Mergers and acquisitions accounting and the diversification discount, *Journal of Finance* (forthcoming).
- Diamond, Peter A., 1993, Search, sticky prices, and inflation, *Review of Economic Studies* 60, 53–68.
- Dixit, Avinash, and Victor Norman, 1980, *Theory of International Trade* (Cambridge University Press).
- Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: the impact of prices on takeovers, *Journal of Finance* 67, 933–971.
- Fan, Joseph, and Larry Lang, 2000, The measurement of relatedness: An application to corporate diversification, *Journal of Business* 73, 629–60.
- Garcia, L.B., and W.I. Zangwill, 1982, *Pathways to Solutions, Fixed Points, and Equilibria* (Prentice-Hall).
- Gomes, Joao, and Dmitry Livdan, 2004, Optimal diversification: reconciling theory and evidence, *Journal of Finance* 59, 505–535.
- Gormley, Todd A., and David A. Matsa, 2013, Common errors: How to (and not to) control for unobserved heterogeneity, *Review of Financial Studies* (Forthcoming) .
- Graham, John. R., Michael L. Lemmon, and Jack G. Wolf, 2002, Does corporate diversification destroy value?, *Journal of Finance* 57, 695–719.
- Hackbarth, Dirk, and Erwan Morellec, 2008, Stock returns in mergers and acquisitions, *Journal of Finance* 63, 1213–1252.
- Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.
- Jackson, Matthew O., 2008, *Social and Economic Networks* (Princeton University Press).

- Kaplan, S., and M. S. Weisbach, 1992, The success of acquisitions: evidence from divestitures, *Journal of Finance* 47, 107–138.
- Maksimovic, Vojislav, and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently?, *Journal of Finance* 57, 721–767.
- Maskin, Eric, and Jean Tirole, 2001, Markov perfect equilibrium I: Observable actions, *Journal of Economic Theory* 100, 191–219.
- Matsusaka, John G., 2001, Corporate diversification, value maximization and organizational capabilities, *Journal of Business* 74, 409–431.
- Mortensen, Dale T., and Christopher A. Pissarides, 1994, Job creation and job destruction in the theory of unemployment, *Review of Economic Studies* 61, 397–415.
- Obreja, Iulian, and Chris Telmer, 2013, Accounting for low-frequency variation in Tobin's q, Working paper.
- Rajan, Raghuram G., Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance* 55, 35–79.
- Ravenscraft, D. J., and F. M. Scherer, 1987, *Mergers, sell-offs, and economic efficiency* (Brookings Institution).
- Rhodes-Kropf, Matthew, and David T. Robinson, 2008, The market for mergers and the boundaries of the firm, *Journal of Finance* 63, 1169–1211.
- Ricardo, David, 1817, *The Principle of Political Economy and Taxation* (Gearney Press (1973)).
- Sanzhar, Sergey V., 2006, Discounted but not diversified: Organizational structure and conglomerate discount, Working Paper, available at SSRN.
- Scharfstein, David S., Robert Gertner, and Eric Powers, 2002, Learning about internal capital markets from corporate spinoffs, *Journal of Finance* 57, 2479–2506.

Scharfstein, David S., and Jeremy C. Stein, 2000, The dark side of internal capital markets: divisional rent-seeking and inefficient investment, *Journal of Finance* 55, 2537–2564.

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

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

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

Appendix

TABLE OF CONTENTS

- A.1. Construction of segment-distance variable
- A.2. Proofs
- A.3. Details about calibration with time-varying σ
- A.4. Extension: model with truncated matching
- A.5. Summary statistics and variable definitions

A.1 Construction of segment-distance variable

We adopt the approach in Anjos and Fracassi (2014), who use input-output flows to construct an industry-network representation of the U.S. economy. Conglomerate segment distance is defined formally as follows:

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

where \mathcal{I} denotes the set of industries a diversified firm participates in, M is the size of this set, and l_{ij} the length of the shortest path between industries i and j . This shortest path is computed by considering the overall industry network of the economy. We further scale this measure by its unconditional mean.

Our network builds on the benchmark input-output table for the year 1997 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.^{A.1} The industry and commodity flows are aggregated into 470 industries, a similar level of aggregation as the 4-digit SIC code. We use such industry classification, rather than more conventional classifications such as SIC or NAICS, because the input-output tables reporting the flow of goods and services between industries come from the Bureau of Economic Analysis. Detailed input-output tables are prepared by the BEA every 5 years.

Next we detail the computation of the shortest paths l_{ij} . First we create a square matrix of flows. We use flows from the USE tables, which report a dollar flow from commodity i to industry j , and where each industry has an assigned primary commodity; we denote this flow by f_{ij} . We normalize these flows by creating a transformed flow variable $\bar{f}_{i,j}$:

$$\bar{f}_{i,j} := \frac{0.5 (f_{ij} + f_{ji})}{0.25 \left(\sum_i f_{ij} + \sum_j f_{ij} + \sum_i f_{ji} + \sum_j f_{ji} \right)}. \quad (\text{A.2})$$

^{A.1}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, 2014).

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 an industry pair, by taking the inverse of the normalized flow:

$$d_{ij} = \frac{1}{f_{ij}} \tag{A.3}$$

With the adjacent distances we can now construct an industry network, which is a weighted undirected graph. Given the industry network, we compute the weighted shortest path (one can think of distance as a cost) between any two industries, l_{ij} , by determining the total distance of the optimal path (i.e. the one that minimizes total distance or cost).^{A.2}

A.2 Proofs

Proof of proposition 1.

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

Case 1: $z \leq \sigma$

Consider the left circle in figure A.1. Let us denote the six adjacent regions in the following way. Starting at 0 and going clockwise until z defines region \mathcal{R}_1 ; starting at z and going clockwise until σ defines region \mathcal{R}_2 ; and so forth. The location of the project generated by i can occur in regions 1, 2, 5, or 6; the location of the project generated by j can occur in regions 1, 2, 3, or 6. Since profits are linear in distance between BUs and projects, the optimal allocation 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.

^{A.2}These network measures were computed using MATLAB BGL routines (available at <http://www.mathworks.nl/matlabcentral/fileexchange/10922>), namely the dijkstra algorithm for minimal travel costs.

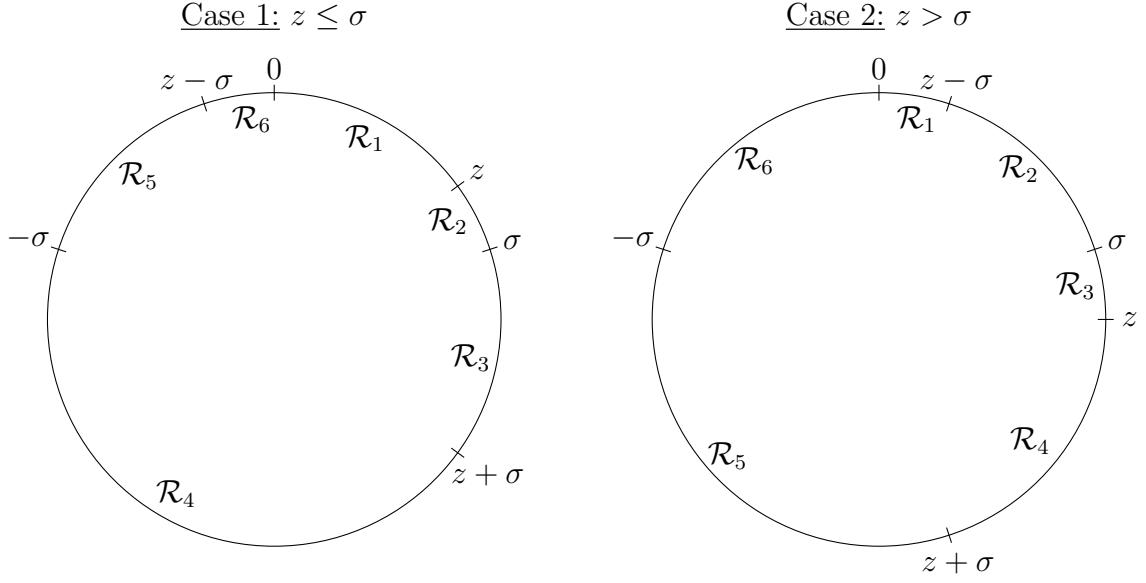


Figure A.1: Splitting the circle into regions. In the left example, $\sigma = 0.2$ and $z = 0.15$. In the right example, $\sigma = 0.2$ and $z = 0.25$.

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

$$\begin{aligned}
E[z_{i,P_i^*}] &= \\
&= \Pr\{\alpha_{P_i} \in \mathcal{R}_1\} \left[\Pr\{\alpha_{P_j} \in \mathcal{R}_1\} E[\min(z_{i,P_i}, z_{i,P_j}) | \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_1] + \right. \\
&\quad \left. + \Pr\{\alpha_{P_j} \in \mathcal{R}_6\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_6] + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}) E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_1] \right] + \\
&\quad + \Pr\{\alpha_{P_i} \in \mathcal{R}_2\} \left[\Pr\{\alpha_{P_j} \in \mathcal{R}_1\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_1] + \right. \\
&\quad \left. + \Pr\{\alpha_{P_j} \in \mathcal{R}_6\} E[z_{i,P_j} | \alpha_{P_j} \in \mathcal{R}_6] + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_1 \cup \mathcal{R}_6\}) E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_2] \right] + \\
&\quad + \Pr\{\alpha_{P_i} \in \mathcal{R}_5\} E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_5] + \Pr\{\alpha_{P_i} \in \mathcal{R}_6\} E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_6]. \tag{A.4}
\end{aligned}$$

The expression (as a function of parameters) of each of the components in equation (A.4) is presented in table A.2.

We are omitting the explicit integration procedures, since all conditional distributions are uniform (in the relevant region), so probabilities and expected distances are generally sim-

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 $z \leq \sigma$.

ple functions of (region) arc length; the slightly more complex case is the computation of $E[\min(z_{i,P_i}, z_{j,P_j})|\dots]$, where we used a standard result on order statistics for random variables drawn from independent uniform distributions.^{A.3}

Inserting the expressions from table A.2 into equation (A.4), and after a few steps of algebra, one obtains

$$E[z_{i,P_i^*}] = \frac{1}{24\sigma^2} (-z^3 + 6\sigma z^2 - 6\sigma^2 z + 12\sigma^3), \quad (\text{A.5})$$

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

^{A.3}The expected value of the k -th order statistic for a sequence of n independent uniform random variables on the unit interval is given by

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

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

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

Table A.2: Auxiliary table for derivation of equation (A.5).

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 region 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

$$\begin{aligned}
E[z_{i,P_i^*}] &= \\
&= \Pr\{\alpha_{P_i} \in \mathcal{R}_1\}E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_1] + \Pr\{\alpha_{P_i} \in \mathcal{R}_6\}E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_6] \\
&+ \Pr\{\alpha_{P_i} \in \mathcal{R}_2\} \left[\Pr\{\alpha_{P_j} \in \mathcal{R}_2\}E[\min(z_{i,P_i}, z_{i,P_j}) | \alpha_{P_i}, \alpha_{P_j} \in \mathcal{R}_2] + \right. \\
&\left. + (1 - \Pr\{\alpha_{P_j} \in \mathcal{R}_2\})E[z_{i,P_i} | \alpha_{P_i} \in \mathcal{R}_2] \right]. \tag{A.6}
\end{aligned}$$

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 $z > \sigma$.

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

Table A.4: Auxiliary table for derivation of equation (A.7).

The expression of each of the components in equation (A.6) is presented in table A.4.

Inserting the expressions from table A.4 into equation (A.6), and after a few steps of algebra, one obtains

$$E[z_{i,P_i^*}] = \frac{1}{24\sigma^2} (z^3 - 6\sigma z^2 + 12\sigma^2 z + 4\sigma^3), \quad (\text{A.7})$$

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.

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

We focus on the equilibrium where mergers take place (the other case is trivial). The solution to the firm's optimization problem (6) 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 (6), given financial markets' equilibrium, needs to verify the following conditions (where for notational

simplicity we set $\tau = 0$):

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

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

$$\begin{cases} rJ_2(z, t) = [\pi_1(z) - \beta] + \lambda_2[J_0 - J_2(z)] & \text{(A.8)} \\ rJ_1(z, t) = \pi_1(z) + \lambda_1[J_2(z) - J_1(z)] & \text{(A.9)} \\ rJ_0 = \pi_0 + \{E[J_1(z, t + dt)|z \in [z, \bar{z}]] - J_0\}. & \text{(A.10)} \end{cases}$$

Manipulation of equations (A.8)-(A.9) straightforwardly yields expressions (7)-(8) in the proposition. Using equation (8), we can write

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

as

$$\int_{z_L}^{z_H} \left(\frac{1}{z_H - z_L} \right) J_1(z) dz = \frac{\bar{\pi}_1(r + \lambda_1 + \lambda_2) - \lambda_1\beta + \lambda_1\lambda_2J_0}{(r + \lambda_1)(r + \lambda_2)}.$$

Inserting the above expression into equation (A.10), and solving for J_0 , one obtains equation (9) 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 (13) in the proposition. The first result obtains along similar lines. The mass of single-segment firms becoming diversified over an infinitesimal dt , $p\lambda_0q dt$, 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 (12). 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:

$$J_1(z^*) \leq \frac{\pi_0}{r} \Leftrightarrow \frac{\pi_1(z^*)(r + \lambda_1 + \lambda_2) - \lambda_1\beta + \lambda_1\lambda_2J_0}{(r + \lambda_1)(r + \lambda_2)} \leq \frac{\pi_0}{r},$$

where we used equation (8). 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 $\phi\sigma \leq C$. 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 $\phi\sigma \leq C$, 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 non-zero $\{z_L, z_H\}$ 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 $\phi\sigma \leq C$ (see for example Garcia and Zangwill, 1982 for more technical details). ■

A.3 Details about calibration with time-varying σ

A.3.1 Solution method

With time-varying σ firms still face the optimization problem described in (6), except now merger-acceptance policies are time-varying. The only caveat to the similarity in optimiza-

tion 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, we assume that only bad conglomerates can refocus. We do this mainly for technical reasons, since the solution to the unconstrained case is quite more complicated. Furthermore, we do not expect this effect to change our magnitudes importantly.

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). In particular, we set this magnitude at 0.22.
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 0.14, which implies an output growth rate of 1.8% for single-segment firms, within the relevant time period of 1990-2011). 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 (8) and (7).
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):

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

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 $c_g(z, t)$:

$$\begin{aligned}
\frac{\Delta p(t)}{\delta_t} &= \int_0^{1/2} c_b(z, t - 1)\lambda_2 dz - p(t - 1)\lambda_0 q(t) \\
\frac{\Delta c_g(z, t)}{\delta_t} &= p(t - 1)\lambda_0 q(t) dz - c_g(z, t - 1)\lambda_1 \\
\frac{\Delta c_b(z, t)}{\delta_t} &= c_g(z, t - 1)\lambda_1 - c_b(z, t - 1)\lambda_2
\end{aligned}$$

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 1997-2011 interval.

A.3.2 Additional outputs

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

Table A.5: Model outputs and data: steady-state vs. no-stationary model. The table shows key moments, both in the steady-state (SS) calibration and the non-stationary (NS) calibration (averages across periods). “Single-Seg. Value” is the Tobin’s Q of single-segment firms; “Prop. Single-Seg.” is the proportion of assets in the economy allocated to single-segment firms; “Av. Excess Value” is the unconditional excess value of conglomerates; “Merger-Acceptance Prob. (q)” stands for the likelihood that a single-segment BU presented with a merger opportunity will accept it; and “Av. Div. Returns” stands for the average announcement returns of diversifying mergers.

Moment	SS-Calibration	NS-Calibration
Single-Seg. Value	1.53	1.77
Prop. Single-Seg.	50%	49%
Av. Excess Value	-0.24	-0.22
Merger-Acceptance Prob. (q)	16%	15%
Av. Div. Returns	3.5%	2.8%

The main differences are a higher value of single-segment firms J_0 , as well as lower average diversifying-merger returns. 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 (note that the average normalized dollar amount is similar: $2.8\% \times 1.77 = 0.050$, and $3.5\% \times 1.53 = 0.054$).

Figure A.2 plots the main outputs of the model, but for the whole simulation period (150 years).

Some differences arise with respect to the narrow 22-year period shown in figure 7 in the main text. First, the bottom-right panel shows that excess value evolves non-monotonically, and in particular decreases for some periods around the 70-year mark. This effect is due to the fact that after a certain period, mergers simply cease (see red dashed line in top-left panel), which means that there is no entry of “fresh” good conglomerates. As time goes by, existing good conglomerates eventually turn bad (see dotted black line in top-left panel), making average excess value decrease. Second, average segment distance converges to a constant. This constant is determined by the policies associated with the last diversifying mergers that take place in the economy, which show up in the vanishing population of conglomerates.

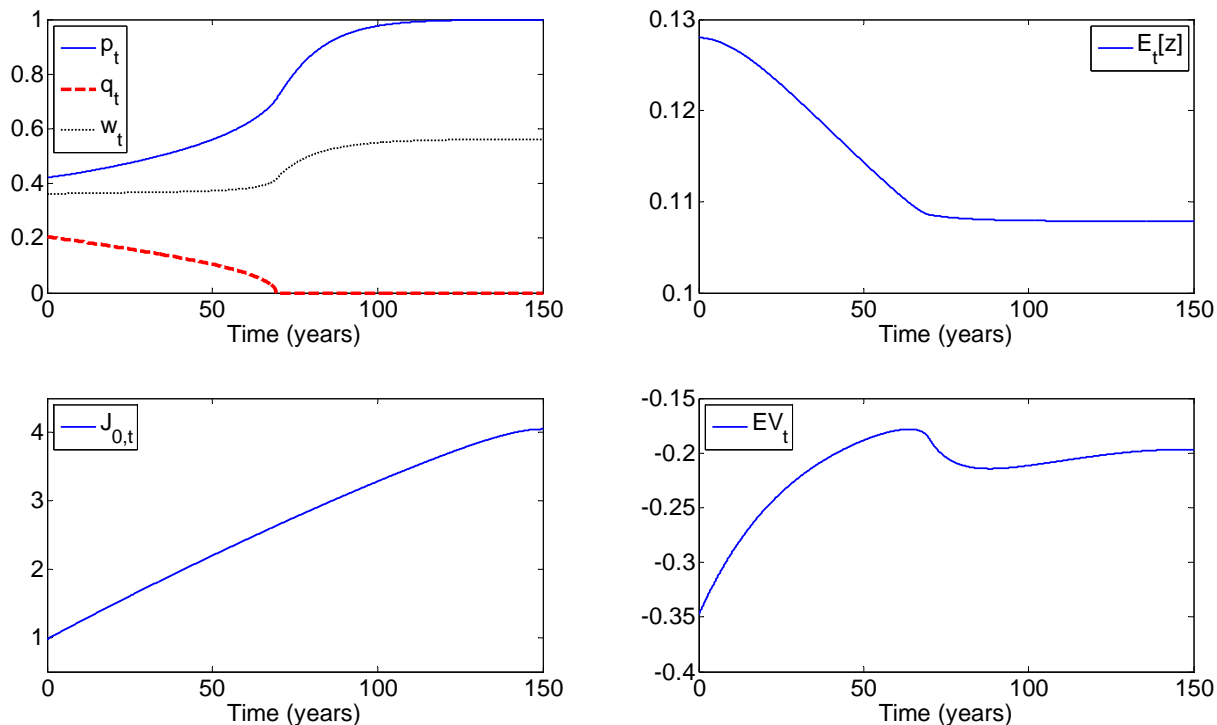


Figure A.2: Calibration with Time-Varying Specialization: Key Outputs (Long Time Horizon). The top-left panel shows three magnitudes: (i) the proportion of single-segment assets in the economy (p), (ii) the probability that a merger opportunity is carried out (q), and the fraction of bad conglomerates in the economy (w); the top-right panel shows the average diversified-firm segment distance; the bottom-left panel plots value of single-segment firms; and the bottom-right panel plots conglomerate excess value.

A.3.3 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.3 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 22-year period, which in data corresponds to the interval 1997-2011, 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.

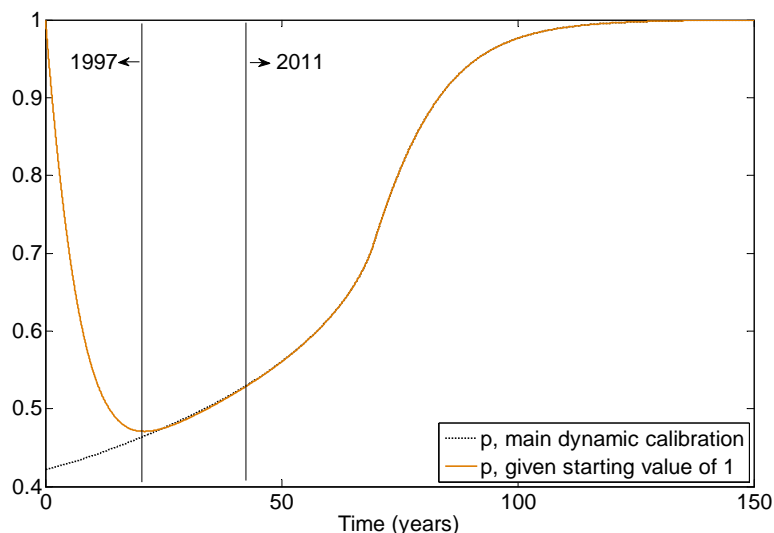


Figure A.3: Initial conditions: robustness check. The figure plots the evolution of p under alternative initial conditions: 99.9% of all firms are single-segment at time 0; and 99.9% of all conglomerates are good at time 0.

A.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, and thus have a support that is proportional to σ . We define this truncation in the simplest possible way, requiring that matches occur uniformly in the interval $[0, \eta\sigma]$. When this new constraint is binding, we are able to match the cross-sectional empirical pattern presented in section 5, namely that excess value increases in segment distance. For the extended model, we replicate the calibration steps of the main model: first we use a steady-state calibration to pin down most parameters; second we introduce time variation in σ (same choice as the one describe in section A.3). In order to identify the new parameter η we choose the difference in excess value across high- and low-segment-distance conglomerates,

$$\Delta EV := EV|_{z > median} - EV|_{z \leq median} \quad (\text{A.11})$$

which in our data is about 0.06.

Table A.6 summarizes the choice of parameters. Table A.7 reports key levels (compares to table 2 for the main model). Table A.8 reports key trends (compares to table 3 for the main

Table A.6: Calibrated parameters. The table shows the magnitude of each model parameter used in the extended-model calibration.

Parameter	Value
r	0.10
η	0.45
λ_0	0.21
λ_1	0.09
λ_2	0.20
β	0.41
ϕ	8.70
σ	0.20

Table A.7: Model outputs and data: truncated matches (1/2). The table shows key moments, both in the calibration and in data; for the steady-state calibration. “Single-Seg. Value” is the Tobin’s Q of single-segment firms; “Prop. Single-Seg.” is the proportion of assets in the economy allocated to single-segment firms; “Av. Excess Value” is the unconditional excess value of conglomerates; “ Δ Excess Value” is the difference in excess value between above-median-segment-distance and below-median-segment-distance conglomerates; “Probab. of M&A” stands for the likelihood that a single-segment BU engaged in at least one merger deal; and “Av. Div. Returns” stands for the average announcement returns of diversifying mergers.

Moment	Model Counterpart	Calibration Output	Data/target
Single-Seg. Value	J_0	1.34	2.00
Prop. Single-Seg.	p	52%	55%
Av. Excess Value	$\frac{wE[J_2]+(1-w)E[J_1]-J_0}{J_0}$	-0.21	-0.28
Δ Excess Value	$\frac{E[J z>z_{median}]-E[J z\leq z_{median}]}{J_0}$	0.04	0.06
Probab. of M&A	$1 - e^{-\lambda_0 q}$	5.6%	6.0%
Av. Div. Returns	$\frac{E[J_1]-J_0}{J_0}$	4.6%	3.8%

model). The truncated model can fit data well, and in particular explains two-thirds of the relation between segment distance and excess value (Δ EV is 0.04 in the model and 0.06 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, $\lambda_0 = 0.21$, whereas $\lambda_0 = 0.37$ 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.

Table A.8: Model outputs and data: truncated matches (2/2). The table compares the annual average growth rates implied by the model for each variable, and compares it to a target interval in data. J_0 is the value of single-segment firms, $|EV|$ is absolute average excess value, p is the fraction of single-segment firms, and \bar{z} is average segment distance.

Variable	Model-implied growth rate	Data target
p	0.43%	[0%, 3%]
\bar{z}	-0.14%	[-0.9%, -0.4%]
J_0	1.59%	[1%, 5%]
$ EV $	-1.21%	[-5%, 2%]

A.5 Summary statistics and variable definitions

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

- *Excess EBIT/Sales*: The difference between the EBIT/sales of a conglomerate and the EBIT/sales of a similar portfolio of single-segment firms. We did not take the log difference as in other excess measures because in many cases EBIT/Sales is negative (Source: COMPUSTAT Segment and Authors Calculations).
- *Excess Value*: The log-difference between the Tobin's Q of a conglomerate and the assets-weighted Tobin's Q of a similar portfolio of single-segment firms, using the detailed Input-Output industry classification system (Source: CRSP, COMPUSTAT, BEA, and Authors Calculations).
- *Number of Segments*: The number of unique segments of a conglomerate using the detailed Input-Output industry classification system (Source: COMPUSTAT SEGMENTS and BEA).
- *Related Segments*: The number of unique segments of a conglomerate using the detailed Input-Output industry classification system, minus the number of unique segments of a conglomerate using the 3-digit Input-Output industry classification system, following Berger and Ofek (1995) (Source: COMPUSTAT SEGMENTS and BEA).
- *Sales*: Gross sales reduced by cash discounts, trade discounts, and returned sales (Source: SALE variable in COMPUSTAT).
- *Segment Distance*: the distance between any two industries the conglomerate participates in, averaged across all pairs (Source: COMPUSTAT SEGMENTS, BEA, and Authors Calculations). We scale the raw variable by its unconditional mean.
- *Tobin's Q* : The sum of total assets (AT) minus the book value of equity (BE) plus the market capitalization (Stock Price at the end of the year (PRCC_F) times the number of shares outstanding (CSHO)), divided by the total assets (AT) (Source: COMPUSTAT).

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

Table A.9: Summary Statistics. The table presents summary statistics for each variable.

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

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