Survival and Conglomerate Value

Finance Working Paper N° 600/2019
July 2022

Michela Altieri
Luiss Guido Carli University

Giovanna Nicodano
University of Torino, Collegio Carlo Alberto, Netspar and ECGI

ECGI Working Paper Series in Finance

Survival and Conglomerate Value

Working Paper N° 600/2019
July 2022

Michela Altieri
Giovanna Nicodano

We gratefully acknowledge useful suggestions by Bernardo Bortolotti, Matthijs Breugem, Gilles Chemla, Raffaele Corvino, Claudia Custodio, Marc Deloofs, Bernard Dumas, Ester Faia, Alex Guembel, Marco Da Rin, Alberto Manconi, Roberto Marf˝e, Arie Melnik, Loriana Pelizzon, Luca Regis, Alessandro Sembenelli, Bel´en Villalonga, Luofu Ye, participants at the 15th Corporate Finance Day in Antwerp, the SIE Conference, the seminars at Collegio Carlo Alberto, Imperial College London, SAFE Frankfurt, the Toulouse School of Economics, the University of Bolzano and of Venezia, and the VU Finance Department. Giovanna Nicodano is grateful to the MIUR initiative "Dipartimenti di Eccellenza." Francesco Massazza and Andrea Romeo provided excellent research assistance.

Abstract

This paper shows how a selection bias generates an apparent value reduction for firms that survive longer in the sample. We test our model on conglomerate firms and find that this bias reduces the value of surviving conglomerates with lower default probability relative to those with higher default probability. The data also indicate that only conglomerates closest to distress have the same value as their focused peers. We show that the selection bias disappears once we control for firm differential ability to survive in the sample. We thus shed new light on the effect of survival-enhancing coinsurance on firm value.

Keywords: survivorship bias, default, coinsurance, contagion, conglomerate discount

JEL Classifications: G1, G14, G3

Michela Altieri*
Assistant Professor of Finance
Luiss Guido Carli University
viale Romania 32
Rome , ND RM 00197, Italy
phone: + 0685225458
e-mail: m.altieri@luiss.it

Giovanna Nicodano
Professor of Financial Economics
University of Torino, Collegio Carlo Alberto and Esomas
Corso Unione Sovietica 218/bis
10134 Torino, Italy
phone:
e-mail: giovanna.nicodano@unito.it

*Corresponding Author

Survival and Conglomerate Value

Michela Altieri* Giovanna Nicodano†

June 15, 2022

Abstract

This paper shows how a selection bias generates an apparent value reduction for firms that survive longer in the sample. We test our model on conglomerate firms and find that this bias reduces the value of surviving conglomerates with lower default probability relative to those with higher default probability. The data also indicate that only conglomerates closest to distress have the same value as their focused peers. We show that the selection bias disappears once we control for firm differential ability to survive in the sample. We thus shed new light on the effect of survival-enhancing coinsurance on firm value.

JEL Classification: G1, G14, G3.

Keywords: survivorship bias, default, coinsurance, contagion, conglomerate discount.

*Department of Business and Management, Luiss Guido Carli, Rome. Email: maltieri@luiss.it

†Corresponding Author, Collegio Carlo Alberto and Esomas, University of Torino, CEPR and ECGI. Postal address: Piazza Vincenzo Arbarello, 8, 10122 Torino (TO). Mobile phone: +39 3332183983. Email: giovanna.nicodano@unito.it

We gratefully acknowledge useful suggestions by Bernardo Bortolotti, Matthijs Breugem, Gilles Chemla, Raffaele Corvino, Claudia Custodio, Marc Deloofs, Bernard Dumas, Ester Faia, Alex Guembel, Marco Da Rin, Alberto Manconi, Roberto Marfé, Arie Melnik, Loriana Pelizzon, Luca Regis, Alessandro Sembenelli, Belén Villalonga, Luofu Ye, participants at the 15th Corporate Finance Day in Antwerp, the SIE Conference, the seminars at Collegio Carlo Alberto, Imperial College London, SAFE Frankfurt, the Toulouse School of Economics, the University of Bolzano and of Venezia, and the VU Finance Department. Giovanna Nicodano is grateful to the MIUR initiative "Dipartimenti di Eccellenza." Francesco Massazza and Andrea Romeo provided excellent research assistance. The usual disclaimer applies.

1 Introduction

Conglomerates represent the majority of large US firms and cover a relevant share of the total assets, investments and financing in the economy. Past literature has extensively investigated the positive relationship between expected value and coinsurance in these firms that diversify across multiple industries (Banal-Estañol et al. (2013), Leland (2007), Lewellen (1971)). Such relationship is however missing in the data even if coinsurance does reduce the expected cost of capital (see Hann et al. (2013)). A confounding yet neglected aspect is that some conglomerates survive longer than others in the sample. Thus, the relationship between conglomerate value and survival probability may be different in the survivors' sample and in the population of conglomerates due to a selection (or survivorship) bias.

This paper shows how the survivorship bias affects the sign of the relationship between conglomerate value and survival probability both through a model and in US Compustat data. Specifically, we show that the survivorship bias generates a value discount for firms that survive longer in the sample, when there is no discount for them in the population of firms. This discount appears both within the conglomerate sample, and across conglomerates and equivalent (portfolios of) focused firms. We explain this pattern as follows.

Let conglomerates and focused firms have the same expected values at t_0, when conglomerates are formed. Then, at t_1, the two sets of companies will have the same average price only if they have the same default probability. If instead, conglomerates have lower default probability than focused firms, as assumed in Lewellen (1971), then fewer of them will have disappeared during downturns. Because of this different truncation, the sample of surviving conglomerates will display a lower average price than the sample of their focused peers at time t_1. Thus, the survivorship bias will affect the sign of the value difference (the so called conglomerate discount). This occurs even if they would have had the same average values in case the sample could include all the defaulted firms. The same reasoning however holds within the conglomerate sample, in that the conglomerates with higher survival will appear to have lower values than the ones with lower survival. It follows that conglomerates with higher survival probabilities will also appear to have higher discounts relative to focused companies.

To formalize this intuition, we build a reduced-form model based on Boot and Schmeits (2000), which shows how firms with higher survival probability appear to be more discounted due to sample selection whenever the superior ability to survive is not controlled for. The model also clarifies that the sample selection bias does not affect conglomerate value at t_0, the stage of conglomerate formation, because the selection bias is due to firm disappearance later on.

We bring our hypothesis to the data on a sample of US firms starting from 1980 until 2014. First, the estimates of the default probability of US companies, based on the survival analysis in Campbell et al. (2008), confirm that conglomerates survive on average more often than a similar portfolio of focused firms, in accord with the early evidence in Borghesi et al. (2007). Second, we investigate the relationship between survival probability and the conglomerate discount. Specifically, we introduce a new variable labeled as the "excess-default probability" which measures the excess default probability of a firm relative to its industry peers. Next, we examine the covariation between the excess value and the excess default probability in the cross-section of conglomerates and focused companies.

Consistent with our hypothesis, we find no discount for conglomerates that are closest to distress, while the ones belonging to the top quantile of survival probability trade at a severe discount (10%). Moreover, once we control for the effect of default probability on firms' excess value, we find that the discount disappears in the pool of conglomerate firms. These results survive several refinements. Among them, we control for the age of surviving companies, in order to account for better growth opportunities of smaller and younger firms. We use different measures for conglomeration, we add firm fixed effects to the model and employ alternative metrics of firm survival probability.

Finally, we investigate the discount of newly formed conglomerates. Because we assume that the discount is given by the survival in the sample, we should not observe a different value between those new conglomerates and focused firms. Similarly, we should not see any negative association between the discount and conglomerate survival. To consider the endogenous nature of conglomeration, we apply the longitudinal approach of Lang and Stulz (1994) and Graham et al. (2002). Specifically, we identify a set of focused firms at the time of the switching from single to multi-segment firms.[1] Consistent with both predictions, we show that the *ex-ante* discount is both not statistically different from zero and insensitive to variation in survival probability across newly formed conglomerates. In sum, several distinct implications of the survivorship bias conjecture are not rejected in the data.

Our contribution is two-folds. On the one hand we point out that a selection bias reduces the value of surviving companies that have lower default probability relative to surviving companies with higher default probability. It implies that comparisons of realized sample prices across conglomerates with

[1]We use the terms "conglomerate," "diversified firm" and "multi-segment firm" interchangeably to refer to companies that report multiple operating segments in the 10-K. Similarly, we use the terms "focused firm" "single-segment firm" and "standalone firm" interchangeably to refer to a firm that does not report multiple operating segments. We also use the terms "segment units," "segments," "business units," and "operating units" interchangeably to refer to a business unit with separate financial reporting in the 10-K.

heterogeneous survival do not reflect comparisons of their ex-ante expected values. On the other hand, we show that survival-enhancing conglomeration paradoxically results in a reduction of their value relative to focused peers due to such selection bias. Thus, if the researcher finds no conglomerate discount in a dataset (Villalonga (2004a)) or when accounting for goodwill (Custodio (2014), then our insight implies that the ex-ante value of conglomerates exceeds the one of focused units. We thus provide new light on the covariation between excess survival and excess value of conglomerates due to a survivorship bias, which is not considered by any previous theoretical or empirical framework. Importantly, such bias confounds our assessment of the benefits of coinsurance.

The rest of the paper proceeds as follow. Section 1.1 reviews closely related literature. Section 2 presents our model of expected firm value. Section 3 determines market values and examines the robustness of our results. Section 4 investigates the empirical relation between value and default probability. Conclusions follow. Appendix A reports the details of the variables used in the analysis. Appendix B.1 defines the complete set of variables used in the analysis, along with descriptive statistics (Table B.1), and some extensions of the baseline model. Appendix C reports the estimation of firms' default probability.

1.1 Related Literature

Since the early work of Banz and Breen (1986), scholars have been aware of the ex-post-selection bias. This disturbs the comparison of prices and returns across companies displaying different price/earnings ratios and book-to-market ratios (see for instance Kothari et al. (1995)). This paper shows that the sample selection bias affects relative values of firms in proportion to relative survival probability. It provides and tests a simple model of the survivorship bias, and indicates a method to undo the bias.

Brown et al. (1995) highlight that survival distorts return predictability and the equity premium. In the data the probability of an equity market implosion does not however seem high enough to account for the observed equity premium (Julliard and Ghosh (2012), Li and Xu (2002)). Our paper draws the attention to the survival probability of individual firms, which seems considerably low (Bessembinder, 2018), in order to pin down the confounding effects of the survivorship bias on conglomerate value.

Our model builds on company diversification theories that highlight the coinsurance-contagion trade-off in conglomerates. These theories show that diversification on the one hand allows for coinsurance between conglomerate units exposed to less-than-perfectly-correlated industry shocks (as in Boot and Schmeits (2000) and Lewellen (1971)), thereby increasing survival and ex-ante value. On the other hand, it may result in contagion as unprofitable units may drag profitable ones into bankruptcy (as in Banal-

Estañol et al. (2013) and Leland (2007)), thereby decreasing survival and ex-ante value. Our model departs from previous ones by providing a comprehensive analysis of the opposite effects of such trade-off on the ex-post conglomerate value, that are brought about by sample selection.

For the sake of simplicity, our model sidesteps the analysis of any operational (in)efficiency in conglomerates stemming from the internal capital market (Almeida and Philippon (2007), Kuppuswamy and Villalonga (2016), Matvos et al. (2018), Rajan et al. (2000), Stein (2002), Yan et al. (2010)), employees' incentives (Fulghieri and Sevilir, 2011), or production decisions (Alonso et al., 2015).[2] Our insight would imply that any ex-ante conglomerate discount (or premium) generated by such characteristics is increased (reduced) in proportion to excess conglomerate survival in the sample of survivors'.

This paper also brings the survivorship bias to bear on a known problem in empirical finance, namely the conglomerate discount (see, for example, Lang and Stulz (1994); Berger and Ofek (1995); Whited (2001); Rudolph and Schwetzler (2014); Avramov et al. (2016)). Some prior papers highlight the role of (bankruptcy) risk in generating the conglomerate discount. Lee and Li (2012) find that the lower profitability of conglomerates disappears once taking into account the operational risk. Mansi and Reeb (2002) and Glaser and Mueller (2010) stress that diversification brings about a default risk reduction and an associated transfer from shareholders to bondholders. Thus an (apparent) conglomerate discount arises because of the use of the book value instead of the market value of debt.

While the implications of the survivorship bias may seem identical to those of previous theories of the conglomerate discount, they are not. On the one hand, most previous theories address the size of the discount instead of its covariation with (excess) survival, which is our focus. Furthermore, the survivorship bias theory restricts the relationship between the conglomerate discount and (excess) survival in several ways. On the contrary, each extant theory of the conglomerate discount rationalizes only a subset of them. For instance, endogenous conglomerate diversification stories (such as Campa and Kedia (2002), Villalonga (2004b), Gomes and Livdan (2004), Graham et al. (2002)) imply that the discount is both absent at t_o but present at t_1. However, they do not seem to explain why the discount increases in conglomerate survival at t_1 and does not decrease in survival at t=0. Similarly, the book value conjecture (in both Mansi and Reeb (2002) and Glaser and Mueller (2010) explains why the value of conglomerates, computed using the book value of debt, decreases in survival probability at t_1. However, it does not seem to explain why such relationship is missing at t_0. In conclusion, the survivorship bias story delivers a set of implications that distinguish it from previous theories, shedding new light also on the sources of conglomerate discount.

[2]For the early debate, see Alchian (1969), Williamson (1985), Stein (1997) and Maksimovic and Phillips (2013).

2 The Model

The model, that builds on the setup of Boot and Schmeits (2000), examines a population of firms composed in equal proportion by conglomerates and focused companies. It matches conglomerates, that combine operating units, to a combination of focused companies running only one unit.

In section 2.2 we analyze the *ex-ante* stage, when all companies are alive, assuming no frictions. The (average) price of each type of company at the *ex-ante* stage will then coincide with the respective *ex-ante* expected value. In Proposition 1, we determine the conglomerate discount in the population of companies and its variation with (excess) survival (Proposition 1). This section reproduces, in a simplified set-up, known results in previous literature.

In section 3, we determine the (average) sample price at a later interim stage, after some companies of the two types may have defaulted. The survivorship bias will arise because defaulted companies and their prices are missing from the sample. We use two metrics for the survivorship bias. One is the difference between the average sample price at the interim stage (equal to the expected value conditional on survival), and the average *ex-ante* price (equal to the unconditional expected value), for each company type (Proposition 2). We will show that conglomerates with higher survival probability do not necessarily have higher sample value than riskier conglomerates, as implied by Proposition 1, because of the survivorship bias. The other metric for the survivorship bias is the difference between the average sample market values of the two company types at the interim stage (Proposition 3).

The next section describes the coinsurance-contagion trade-off in conglomerates. Conglomerate diversification affects survival by permitting coinsurance across units, thereby lowering default probability. It may however result in contagion, thereby increasing default probability. This trade-off will determine the risk-adjusted credit spreads of each company type. In order to focus on the differential survival both within and across the two company types, we will rule out other differences such as a unit's profitability, debt needs, and bankruptcy costs, in line with the literature on mergers motivated by purely financial synergies.

2.1 The Coinsurance-Contagion Trade-Off and Survival Probability

Each unit, indexed by $i = (A, B)$, raises an amount of debt D_i to invest in a project at the stage of company creation ($t = 0$). Competitive lenders earn a credit spread R_i, which is determined at $t = 0$ together with the *ex-ante* expected value of each company type. The operating profit of each unit is realized in $t = 1$ and is independently distributed across units. It will be High $\{H\}$ and equal to $X_i > 0$ with probability $p_i \in (0, 1)$, and it will be Low $\{L\}$ and equal to zero with probability $(1 - p_i)$. Our assumptions on the size of operating profits will ensure that each unit generates, in state L, insufficient operating profits to honor its own debt obligations.

At the interim stage, lenders observe a private and perfect signal of future operating profits and may decide to declare bankruptcy. When a company defaults, the (prices of) defaulted companies no longer exist. At this stage, after some companies may have defaulted, we will determine both the average price of survivors and the survivorship bias. When a company defaults, it may have to bear bankruptcy costs. We model these costs as the loss of its future profit conditional on survival, $K_i \geq 0$.

One company type (F) is made of a portfolio of focused companies. Each unit, i, belonging to F is independently liable to its own lenders and has survival probability equal to $p_i^{Sur} = p_i$. Diversified conglomerate (C) combine instead two units and pool their operating profits, so that they are jointly liable *vis-à-vis* lenders. A profitable unit may therefore be able to help the insolvent one, or *vice versa* an unprofitable unit may drag a profitable one into bankruptcy.

To represent this coinsurance-contagion trade-off, we define four states of the world $\{HH, LL, HL, LH\}$ where the first (second) letter in each pair refers to the profit of unit A (B). We let the profit of unit A, in state $\{HL\}$, exceed the combined debt repayment of the two units, whereas the profit of unit B is lower than the combined service of debt.Thus, a conglomerate will default when a lender's signal is either $\{LL\}$ or $\{LH\}$, the latter being a contagion state because A's losses drag B into bankruptcy. The conglomerate will survive when the signal is either $\{HH\}$ or $\{HL\}$, the latter being a coinsurance state because profits from A rescue B. The resulting survival probability of conglomerates, p_C^{Sur}, is equal to p_A because the conglomerate survives if and only if unit A survives. [3]

2.2 The Credit Spread and the Ex-ante Expected Values

In this section, we first determine the credit spread charged to each company type. We then determine the *ex-ante* expected value of companies, before any default occurs, which will serve as a benchmark to show the effect of the survivorship bias.

[3] So far, we are following the setup of Boot and Schmeits (2000) without incentive problems, adding instead the assumption of asymmetric profits. This assumption makes contagion possible, a feature that is prominent in other studies of conglomerate mergers such as Banal-Estañol et al. (2013) and Leland (2007). In Appendix A we also allow unit B to coinsure and contaminate unit A.

Lenders of unit i, $i = A, B$, receive debt repayment in state $\{H\}$ and collect nothing in state $\{L\}$. It follows that the credit spread for unit i, R_i, satisfying the lenders' zero expected profit condition, $(1 - p_i) \times 0 + p_i R_i = D_i$, is equal to:

$$R_i = D_i p_i^{-1} \,. \tag{1}$$

In turn, conglomerate lenders receive the debt repayment in states $\{HH\}$ and $\{HL\}$. They also recover the cash flow X_B in state $\{LH\}$, when unit A drags the profitable unit B into bankruptcy. The credit spread for the conglomerate, when default states are correctly anticipated, is thus equal to:

$$R_C = [D_A + D_B - p_B(1 - p_A)X_B)]p_A^{-1} \,. \tag{2}$$

This spread solves the zero profit condition, which requires lenders' expected repayments to equal the loan provided at t=0, that is, $[p_A p_B + p_A(1-p_B)]R_C + p_B(1-p_A)X_B = D_A + D_B$. Lenders collect the interest payment when either both units are successful, an event that has a probability of $p_A p_B$, or when there is coinsurance, that is unit A is profitable when B is not, which has probability of $p_A(1 - p_B)$. Moreover, they recover profit, X_B, upon the conglomerate default when there is contagion, with a probability of $p_B(1 - p_A)$.

The Lemma in Appendix A states the ranking of credit spreads across company types while making explicit the cash flow restrictions that support our state space and the derivations of Equations (1)-(3). It shows that:

$$R_C < R_A + R_B \,, \tag{3}$$

Conglomerates thus always enjoy better credit conditions than focused companies. This is due in part to coinsurance, which reduces the chances of default, as in Hann et al. (2013). Paradoxically, and differently from Hann et al. (2013), also contagion across units helps reduce the conglomerate credit spread because lenders recover the operating profit of the healthy unit upon default.

Let us now turn to the expected value of the population of companies at $t = 0$, before any default occurs. Let $\pi_i = p_i X_i - D_i$, for i=A,B, denote the expected profit after the service of debt. Recall that $p_i^{Sur} = p_i$ for focused units, and $p_C^{Sur} = p_A$ for conglomerate firms. Furthermore, recall that coinsurance and contagion probabilities are respectively equal to $p_A(1 - p_B)$ and $p_B(1 - p_A)$. In line with the past literature, we find that:

Proposition 1: *At t=0:*

a. Expected value, V, is non-decreasing in survival probability and is equal to:

$$V_F = \pi_A + \pi_B + p_A^{Sur} K_A + p_B^{Sur} K_B \tag{4}$$

$$V_C = \pi_A + \pi_B + p_C^{Sur}(K_A + K_B) \tag{5}$$

for a focused company and a conglomerate, respectively.

b. The conglomerate expected excess value relative to the focused company, $V_C - V_F$, is positive if, and only if, the coinsurance probability exceeds the probability of contagion.

Part (a) of Proposition 1 shows that the *ex-ante* unconditional expected value of diversified and focused companies coincides, as in the intuitive example in the introduction, when there are no bankruptcy costs. When bankruptcy costs are positive, expected value increases in survival probability for all firm types because higher survival probability saves on bankruptcy costs. This result is a toy replica of previous insight from Banal-Estañol et al. (2013), without tax-distortions, and Leland (2007), with tax distortions. Part (b) of Proposition 1 argues that there is a conglomerate discount at the time of conglomerate formation ($t = 0$) only if conglomerates display higher default probability relative to focused firms. This result holds when we overlook operational (dis)synergies, as in this simplified model.

Despite these known theoretical results, there is little if any evidence that coinsurance increases conglomerate value. Interestingly, Hann et al. (2013) are instead able to show that coinsurance reduces the conglomerate cost of capital, as in our Lemma. To do so, they use ex-ante (rather than ex-post) measures. Ideally, we should similarly rely on the expected value of the population of companies that is not usually available in the data. However, in the empirical section we will be able to measure the *ex-ante* value difference between conglomerates and focused companies in an experiment focusing on newly-formed conglomerates. At that time, we expect no differential sample selection bias between focused units that just became conglomerates and focused units that did not. Therefore, the price difference coincides with *ex-ante* value difference. We will also examine whether all high survival firms (both newly-formed conglomerates and focused ones) display positive *ex-ante* excess value relative to low survival ones, as implied by Proposition 1.

3 The Survivorship Bias

We now determine the average market values of companies that survive into the sample at the interim stage. At this time, the sample of listed companies no longer coincides with the population of listed companies at $t = 0$. Prices are therefore equal to the expected values conditional on company survival.

The following proposition summarizes our finding, concerning the relationship between the prices of each company type and their survival probabilities, which are equal to $p_i^{Sur} = p_i$ for focused units and $p_C^{Sur} = p_A$ for conglomerates:

Proposition 2: *Due to sample selection, at the interim stage the average market value, MV_j, of surviving companies of type $j = F, C$, exceeds its ex-ante expected value, V_j, in direct proportion to its survival probability:*

$$V_F = MV_A \times p_A^{Sur} + MV_B \times p_B^{Sur}\,, \tag{6}$$

$$V_C = MV_C \times p_C^{Sur}\,. \tag{7}$$

Proposition 2 states that, for each company type, the sample price exceeds its *ex-ante* expected value in the population due to the sample truncation, which is proportional to default probability. Note that all individual surviving companies are correctly priced.

We can now provide a first measure of the survivorship bias within a given company type. This is the difference between the average market values of type-j survivors and the average *ex-ante* value for the type-j population. Such measure is $MV_j - V_j$, which is equal to $V_j \times (1 - p_j^{(Sur)})/p_j^{(Sur)}$, for $j = A, B, C$. The survivorship bias is thus larger the lower is the type-specific survival probability. One economic rationale for this result is that, for each company type, the share of bad performers that exit the market is larger the lower is the type-specific survival probability.

We now bring this result to bear on the cross-sectional difference in sample market values between conglomerates and focused firms. Proposition 2 directly implies that a conglomerate discount, generated by the survivorship bias, appears in the sample when conglomerates display excess survival relative to focused companies. When the converse is true, a conglomerate contagion premium appears in the data due to a survivorship bias. These sample differences in market values are a measure of the survivorship bias in the cross section of company types. We can therefore state the following proposition:

Proposition 3: *At the interim stage, the survivorship bias implies that:*

a. There is a conglomerate discount in the sample of survivors if, and only if, the conglomerate survival probability exceeds the survival probability of focused units:

$$MV_C - MV_F = \pi_A[(p_C^{Sur})^{-1} - (p_A^{Sur})^{-1}] + \pi_B[(p_C^{Sur})^{-1} - (p_B^{Sur})^{-1}] < 0\,. \tag{8}$$

b. With positive bankruptcy costs, the larger the ex-ante conglomerate premium is, the larger the conglomerate discount in the sample of survivors will be.

Proposition 3a states that a necessary and sufficient condition for observing in the data a conglomerate discount, brought about by the survivorship bias, is a conglomerate excess survival. According to Proposition 1b, the pattern should be opposite in the population, in that excess survival is associated with an *ex-ante* premium. The survivorship bias therefore changes the sign of the relationship between survival probability and excess value.

Proposition 3a also implies a positive correlation between the sample conglomerate discount, due to the survivorship bias, and the conglomerate (excess) survival probability. On the contrary, the correlation implied by Proposition 1 between the ex-ante excess value and the (excess) survival of the conglomerate population is negative. These opposite patterns will help us identify the survival discount generated by the survivorship bias in the data. Importantly, these opposite patterns survive in an extension of the model, presented in the Appendix A, where coinsurance between conglomerates' units is mutual. That is, there is one state of nature where unit B supports A, as in Boot and Schmeits (2000).

Proposition 3b is not going to have an empirical counterpart. However, it clearly indicates how far the survivorship bias may disturb our understanding of the value of survival-enhancing coinsurance.

Summing up, we have modelled a perfect market where prices at all stages reflect their future cash flows. In such perfect markets, we show the existence of a wedge between the expected value of the population of companies and the sample average prices that is brought about by a survivorship bias. This bias derives from a sample truncation: there are indeed no prices for the defaulted companies. To deliver this insight straightforwardly, we have relied on simplifying assumptions concerning the determinants of *ex-ante* values that will be addressed when taking the model to the data, as explained in the next subsection.

3.1 From the Model to the Data

In the next sections, we test for a survivorship bias in the value of conglomerates relying on four implications of the model. The first implication states that, due to the survivorship bias, the excess conglomerate value is negative when the excess default probability is also negative (see equation (8)). In order to examine this implication in the data, we borrow the definition of excess value from the conglomerate discount literature. We also identify conglomerates as multi-segment companies and control for both firm characteristics and firm fixed effects to account for differences in operating profits and bankruptcy costs.[4] The second implication concerns the sensitivity in the value of surviving conglomerates to their own default probability (see equations (6)). Surviving conglomerates with higher default probabilities ought to display higher values relative to those with lower survival probabilities, everything else equal, due to the survivorship bias.[5] The same equations (6) and (7) deliver the third implication, namely that the market value of surviving firms coincides with their ex-ante value when we appropriately control for their survival probability. The fourth implication is that the relationship between the conglomerate excess value and excess default probability disappears or turns positive, as in equation (5), when conglomerates are just born and there is no survivorship bias.

Let us stress again that support for these implications in the data will reveal the presence of a survivorship bias, leaving no confusion with any of the previous conglomerate discount stories.

We will also examine the robustness of our results when eliminating from the sample firm entries and exits motivated by reasons different from bankruptcy, that are not considered in the model. Finally, the above implications hold also when debt is endogenous (for instance when there is a tax-bankruptcy trade-off, as in Leland (2007) and Luciano and Nicodano (2014)) conditional on debt levels. A robustness check will thus investigate the excess conglomerate values across leverage groups.

4 The Survivorship Bias in Conglomerate Value

4.1 Data and Sample

Our sample combines several data sources from the years 1980-2014. Firstly, we retrieve information on multi-segment companies (that is, conglomerates) from Compustat-Historical Segments. Previous

[4] Hennessy and Whited (2007) indicate that the bankruptcy costs for smaller companies are almost double those of larger companies (15% to 8% of capital).

[5] This implication carries over to the sensitivity of excess conglomerate values to their own default probability.

studies associate each conglomerate segment with similar single-segment companies in the same industry in order to compute excess values. We follow a similar approach, applying both the matching and the sample selection as in Lamont and Polk (2002) and Berger and Ofek (1995). We drop firms that have segments in financial services (SIC 6000-6999) and utilities (SIC 4900-4999), firms with total sales below \$20 million, and firms with aggregated firm segments sales above 1% of total firm sales in Compustat. We also drop segments with missing sales and SIC codes; firms operating in other non-economic activities, such as membership organizations (SIC 8600), private households (SIC 8800), or unclassified services (SIC 8900); and all segments that do not have at least five similar single-unit companies in the same industry. Because firms have debt outstanding in the model, we drop all firms with zero leverage. After those modifications, we have a total of 76,389 firm-year observations (for a total of 10,848 companies) from 1980 to 2014, of which 24,605 (32%) are observations from multi-segment companies.

We retrieve information on company default events from three sources. The first source is the Compustat North America database, which indicates if a company was delisted, and provides the motivation for the delisting. We keep only those delistings attributed to bankruptcy filings and liquidations. The second source is CRSP, which also provides information about all public companies delisted due to a distress event. We keep delistings for liquidation (code 04), bankruptcy (code 574), and stocks that were delisted when the price fell below an acceptable level or for insufficient capital (codes 552 and 560, respectively). The third source is the UCLA-LoPucki Bankruptcy Research Database (BRD), which reports bankruptcy filings (both Chapter 7 and Chapter 11) in the United States bankruptcy courts of the major public companies since October 1, 1979.[6] Similar to (Campbell et al., 2008), we also define failure more broadly to include bankruptcies, financially-driven delisting (reported in CRSP), or D (default) ratings issued by a leading credit rating agency. After combining those sources, we have 1,526 default events from 1980 to 2014, which represent 1.82% of total observations and 14% of the firms in the sample.

For robustness purposes, we also retrieve the default probabilities elaborated from the Credit Research Initiative (CRI) of the University of Singapore (RMI-NUS). The CRI probabilities are built on the forward intensity model developed by Duan et al. (2012). This dataset provides the individual companies' PDs for a subsample of 32,258 US public and private companies. We can match 16,205 observations in our sample, for a total of 3,848 companies.[7] Finally, we retrieve firm characteristics from Compustat North

[6] We are grateful to UCLA-LoPucki for free access to their database until 2014. A company is public according to this source if it filed an Annual Report (Form 10-K or Form 10) with the Securities and Exchange Commission in a year ending not less than three years before the filing of the bankruptcy case. A company is major if assets are worth \$100 million or more, measured in 1980 dollars (about \$280 million in 2020 dollars).

[7] Data are available at https://nuscri.org/en/.

America dataset. Specifically, we keep all firms that have information available on their size, leverage, EBITDA, sales, and capital expenditures. We follow Berger and Ofek (1995) and construct the indicator variable "conglomerate" where the firm operates in several industries. For robustness purposes, we also construct a measure of the cash flow correlation across segment units ($CFCORR$), which captures the coinsurance degree of these conglomerate firms. Details of these variables are in the Appendix B.

Table 1 reports the number of active firms, conglomerates, defaults, and failures par year after applying these modifications. Conglomerates represent 30% of active US companies in our sample and 42% of all assets in Compustat. The average yearly number of default events from 1980 to 2014 is 1.6%, consistent with past results (Campbell et al., 2008). The Cumulative Distress column reports the number of cumulative events of failure from the beginning of the sample. For instance, more than half the number of the 2016 firms that were active in 1980 have failed by the year 2000. A raw indicator of potential sample selection is the comparison between the number of active firms surviving into the sample as of 2014 (1,386) and the number of defunct companies over the sample years (1,526).

The table also reports, for each year, the variation in the number of firms due to mergers, new entries (as in Ramey and Shapiro (1998)), and firms that drop from the sample for unspecified reasons (other exits). Overall, the table shows that the sample is subject to huge variation over time for several reasons, in line with Bessembinder (2018), making it challenging to isolate the survivorship bias in conglomerate value without relying on the implications of our model.

4.1.1 Variables and Univariate Statistics

Following the conglomerate discount literature (see, among others, Berger and Ofek (1995) and Villalonga (2004a)), the firm's excess value is computed as the natural logarithm of the ratio between its market value and its imputed value. The imputed value is the average of the market values of the firms' segment units, the latter being computed by multiplying the segments' sales to the median market-to-sales multiplier of the single-segment companies in the same industry as the segment unit. We implement industry matching using the narrower SIC, including at least five single-segment companies.

In a similar way, we construct the variable "excess default probability" as the natural logarithm of the ratio between a company's default probability and its imputed PD at the end of the year. A negative value of this variable captures a higher survival ability of the firm relative to the median single-segment firm in the industry. The variable is estimated in several steps. We first estimate the conditional default probability based on Campbell et al. (2008). From the estimation in column (3) in table C.1, we retrieve the survival odds ratios and compute the probability of default for each company and for each year accordingly. We also compute the imputed survival probability for each segment as the median survival probabilities across all the focused firms in a specific industry (three digits SIC code), weighted by firm sales. We finally calculate the imputed survival probability for each firm and year as the weighted average across segments of the survival probabilities of the firms' segment units.[8]

Table 2, panel A, reports the univariate statistics of the main variables used in the analysis and the differences in characteristics between conglomerates and focused companies. The t-test differences are estimated with an OLS regression, clustered at the firm level. Consistent with past findings (Villalonga, 2004a), the table shows that conglomerates' mean excess value is negative (-6%), indicating that conglomerates' value is lower than that of their focused industry peers (segments for brevity). The table also shows that conglomerates survive (16%) more in the sample, with their excess default probability being 10% to 24% lower than their industry peers. The average segment cash-flow correlation of conglomerate companies is 43%. However, its variation is considerable, ranging from a minimum of -99% to a maximum of 100%, indicating no diversification.As in past results, conglomerates are larger, older and have both greater leverage and dividend ratios, but display both lower investment and lower sales-to-growth ratios.

This descriptive evidence is not inconsistent with the first implication of our model, indicating that surviving conglomerates displaying negative excess values also have negative excess default due to a survivorship bias. This pattern is also evident in figure 1. On the x-axis, it reports the excess average firm value, ranging from -1.386 to 1.386, as in Villalonga (2004a). On the y-axis, it reports the average excess default probability for conglomerates and focused companies. This figure indicates that conglomerate firms with a severe value discount (left side of the distribution) have a much lower excess default probability (thus higher survival skills) than focused firms.

The statistics in table 2, panel B, show patterns of the excess conglomerate value across survival quantiles that are in line with the second implication of the model. In columns (3), (6), and (9), we report the differences in excess value between conglomerates and focused firms for each quantile of the survival probability. The table shows that conglomerates in the higher survival quantile (above 50%) trade at a discount, while the contrary applies to conglomerates with lower survival skills (10% quantile).

[8]Many works show that segments assets are a biased measure when compared to sales. We however compute a similar measure weighted by assets, finding no substantial differences in the results.

4.2 The Survivorship Bias and the Conglomerate Discount

In this section we test the implications of our model after controlling for company characteristics. In the next subsection we provide evidence in support of the first and the second implications of the model. In the following one we will test the third implication of the model.

4.2.1 Parallel Regressions for Excess Value and Excess Default

According to the first implication of our model, the excess value of surviving conglomerates is the mirror image of their excess default probability due to the survivorship bias. Testing this implication is not straightforward as we cannot regress the excess value onto the excess default probability since firm value and survival probabilities are simultaneously determined. They are thus affected by similar covariates.[9]

A first way to deal with this problem is to separately regress firm excess value and firm excess default probability on those covariates. The dummy variable associated with conglomerates should show a negative (or a positive) sign in both regressions, indicating that a conglomerate discount (or premium) appears when there is excess conglomerate survival (default), as in equation (8) of our model.

We thus start our empirical analysis using the covariates suggested in the literature, as follows:

$$\begin{aligned} ExcessValue_{i,t} &= \alpha + \beta Conglomerate_{it-1} + \beta_1 EBITDA_{it-1} + \beta_2 Salesgrowth_{it-1} + \\ &+ \beta_3 Size_{it-1} + \beta_4 CAPEX_{it-1} + \beta_4 Dividends_{it-1} + \varepsilon_t \\ ExcessDefaultProb_{i,t} &= \alpha + \beta Conglomerate_{it-1} + \beta_1 EBITDA_{it-1} + \beta_2 Salesgrowth_{it-1} + \\ &+ \beta_3 Size_{it-1} + \beta_4 CAPEX_{it-1} + \beta_5 Dividends_{it-1} + \beta_6 NITA_{it-1} + \beta_6 CALC_{it-1} \varepsilon_t \end{aligned} \tag{9}$$

where the vector of controls includes industry (Fama-french 17) and year fixed effects (as in Villalonga (2004a) and many others). Following Borghesi et al. (2007), we also estimate an augmented model

[9] Resorting to a matched sample of conglomerates and stand-alone firms (similar to the approach in Villalonga (2004a)) based on the default probability is not helpful in our setup. Indeed, the covariates that are influenced by the treatment can cause the ignorability assumption to be violated, leading to larger biases when they are included in a matching estimator as controls. In other terms, using a matched sample on leverage, size, age, profitability and, more importantly, the default probability will generate biased results if being a conglomerate affects these characteristics (Rosenbaum (1984)). Wooldridge (2009) also shows that the same applies including the instrumental variables as covariates in a matching estimator, when the treatment is endogenous.

including firm age among the covariates, to control for the life cycle of firms' growth options.[10] As for the excess default probability model, we also control for the current assets/liabilities ratio and the net income over total assets. In all specifications, we cluster at the company level.

Table 3 reports the results. As in previous literature, there is a conglomerate discount equal to 12% in the baseline specification for the excess value in column (1), which drops to 9% in column (2) where controls include company age. Consistent with the survivorship bias hypothesis, the mirror image of this discount is a conglomerate-specific negative excess default probability in columns (3) and (4) of table 3, where we report regressions results for the excess default probability. The estimates show that the default probability of conglomerate firmsconglomerate is firm is, on average, 7.2% lower than the default probability of focused industry peers. For robustness, we repeat the estimation after dropping merged firms, new entries, and exits for reasons different from default (see columns (5) and (6)).[11] In the reduced sample, the discount increases from 9.4% to 10% with excess survival reaching 14%, indicating that the survivorship bias associated with distress is larger than the one arising from all the inflows into and outflows from the stock market during these sample years.

We now turn to the second implication of our model, exploring it through a quantile regression approach (as developed by Koenker and Bassett (1978)) that is robust to outliers. We thus examine the distribution of the excess value conditional on survival probability. We regress firm excess value on four sub-samples divided according to 10^{th}, 25^{th}, 50^{th}, and 100^{th} percentiles of company survival probability, as in equation (9), expecting the conglomerate discount to increase (that is, coefficient of the conglomerate dummy to increase in absolute value) as survival probability increases due to the larger survivorship bias. We also test for the equality of the conglomerate dummy's coefficient between 10^{th} and 50^{th} sub-samples.

Table 4 reports the results of these regressions. Consistent with the second implication of our model, conglomerates show a severe discount (10% - column (4)) in the highest survival probability quantile, while there is none value in the bottom quantile (as in column (1)).The t-test rejects the hypothesis of equal coefficients. As in Proposition 2 of our model, conglomerates display higher value the lower is their survival probability as they are hit by a smaller survivorship bias.

[10]The age of living companies cannot account for the large set of defaults. In the CRSP database, the median time a common stock stays listed is seven and a half years (Bessembinder, 2018).

[11]These changes may affect the survivorship bias. For instance, the acquisition of a distressed company may not only reduce the *ex-ante* expected value of conglomerates (as in Gomes and Livdan (2004) and Graham et al. (2002)) but also contribute to the survivorship bias since low-valuation single-segment companies disappear from the database.

4.2.2 Instrumenting Excess Default Probability

We now turn to the third implication of our model. The excess value we observe in the sample of surviving firms is affected by the survivorship bias, which should vanish when properly taking into account excess default probability.

Since we cannot regress current excess value on current excess default probability due to their joint determination, we instrument the latter with lagged excess default probability (two years). Lagged excess default probability is a valid instrument since it does not affect current excess values, which depend on expected cash flows. At the same time, past default probability is highly correlated with the current one: as the data show, the current excess default probability has 75% correlation with its first lag and 63% with its second lag.[12] In the spirit of Angrist et al. (1996), we regress firm excess values on the lagged excess default probability, and use the regression residual as a measure of the excess value that is free from the survivorship bias, following the model:

$$\begin{cases} x_{1i,t} = \beta x2 | Z = x_{2i,t-2} \\ res_{i,t} = x_{1i,t} - \hat{\beta} x_{2i,t-1} \end{cases} \tag{10}$$

where $x_{1i,t}$ is the firm excess value at time t, and the variable $x_{2i,t-2}$ is the two-periods lag of the excess default probability of the firm. If the survivorship bias is responsible for (a portion of) the conglomerate discount in the sample of surviving firms, we should observe the disappearance of (a reduction in) the conglomerate discount in the second regression. We estimate the firm discount net of the survivorship bias as follows:

$$\begin{aligned} residuals(EV)_{i,t} = \alpha + \beta 1 Conglomerate_{it-1} + \beta_2 EBITDA_{it-1} + \beta_3 Salesgrowth_{it-1} + \\ + \beta_4 Size_{it-1} + \beta_5 CAPEX_{it-1} + \beta_{46} Dividends_{it-1} + \varepsilon_t \end{aligned} \tag{11}$$

We report the results in table 5, columns (1)-(2). In columns (3)-(6) we estimate augmented models using three and four lags for excess default probability. In columns (7)-(9) of table 5, we estimate for robustness a similar specification after dropping all firms with corporate events (new firms, mergers, exits for unspecified reasons). The coefficient of the conglomerate dummy is negligible and often not

[12] To be a valid instrument, lagged default probability has to satisfy two conditions: i) it must be relevant, thus correlated with the endogenous variable - therefore past excess default probability is correlated with the current one; ii) it must have no direct causal effect on the response variable - current stock prices.

statistically different from zero, confirming the contraction of the survivorship bias when controlling for differential survival probability.

Further, we adjust the approach of table 4, regressing the excess value residuals on four sub-samples divided according to 10^{th}, 25^{th}, and 50^{th}, and 100^{th} percentiles of company survival probability. Since the dependent variable is now the excess value, free from any survivorship bias, we expect it not to fall in survival probability, contrary to results in table 4 and in line with equations 5 and 4. The results, in table 6, show a conglomerate discount in the lowest quantile (lowest survival probability) and a conglomerate premium in the upper quantile of survival probability.

These results confirm the key insight of the model, namely the presence of a survivorship bias that adversely affects the value of firms with higher survival probability in the survivors' sample. One consequences is that the conglomerate discount disappears, and possibly turns into a premium, only after accounting for this selection problem. Otherwise, the selection bias draws us to believe that conglomerates have lower value with respect to their industry peers, which survive less in the sample.

This set of results represents the solution to a puzzle. Hann et al. (2013) show that diversified firms have, on average, a lower *ex-ante* cost of capital than comparable portfolios of focused firms, thanks to the coinsurance premium (Lewellen, 1971), which is also predicted by our model in equation 3. However, they cannot reconcile the coexistence of a lower cost of capital and a lower value of conglomerate firms. They suggest that realized returns contain noise, while our results point to a survivorship bias. In the robustness section, we employ additional tests to further corroborate these findings.

4.3 Robustness Tests

We provide several additional tests to confirm our baseline findings. First, we replicate the IV approach adding firm fixed effects, before and after the correction, to control for unobservable characteristics driving both firm ability to survive and firm value. The results, in tables 7, confirm that the 9%-10% conglomerate discount, that is still present despite the firm fixed effects, disappears once we get rid of the survivorship bias by controlling for the excess default probability. In table 8, we also report the estimation of our baseline model with different measures of diversification.

Second, we know that coinsurance benefits may prompt the conglomerate to increase leverage well beyond its focused counterparts (see Leland (2007) and Luciano and Nicodano (2014)), increasing default probability. The relationship between value and default probability differences should then hold conditional on leverage differences. For this reason, we estimate quantile regressions of the company discount where the dependent variable is the excess value, and the samples are divided according to 10^{th}, 25^{th} and 50^{th} percentiles of company leverage, before and after the correction for the survivor probability of firms. Table 9 shows that, consistent with the second implication of our model, the conglomerate discount falls from 13% (column (1)) to 6.5% (column (4)) when leverage increases, bringing the company closer to distress. Once controlling for firm ability to surviving as in equation 11 (columns (5)-(8) of table 9), the discount induced by he selection bias disappears and turns into a premium which is decreasing in leverage (and therefore default probability), as expected from our model.

4.3.1 Ex-ante Value or Ex-post Bias?

In this section, we turn to the fourth implication of our model. The relationship between the conglomerate excess value and default probability should either disappear or turn positive, as in equation (5), when conglomerates are just born. We proceed to measure the conglomerate discount on a sample of newly-established conglomerates whose value is unlikely to be affected by a survivorship bias within a short time span such as one year. To assess the existence of an *ex-ante* discount driven by different characteristics other than survival, we rely on a method originally devised to address the concern that both conglomerate formation and the *ex-ante* discount are jointly endogenous. This method, used by Lang and Stulz (1994), Graham et al. (2002), and Villalonga (2004a), applies a longitudinal approach to the conglomerate discount estimation. Their idea is that, for the discount to be interpreted as evidence of value destruction, the cross-sectional evidence of a discount is insufficient, and one needs to look at changes in the diversification status. In their experiments, firms that switch from focused to conglomerate, accounting for their propensity to diversify, should display an ex-ante discount if diversification is expected to decrease ex-ante value.

We adapt this experiment to determine whether newly-formed conglomerates display a contagion premium (like the one appearing in table 4) in the year after their formation, that is whether firms that become conglomerates and display low survival have higher valuations. We add to the baseline model three variables, which capture the differential ability to survive, and we interact these with our main treated variables, that is, firms that switch from focused to conglomerates. If there are other effects other than the survivorship bias affecting firm value, we should expect that the interaction term being significant at the beginning of the period already.

We begin the experiment by identifying 381 firms that transitioned from being a focused firm to a conglomerate firm.[13] We also restrict our sample to those firms and to focused firms that never change their status. The subsample includes the 381 diversifying firms with data from one year before until one year after diversification plus the 30,173 single-segment firm-years with data one year before and after the change. We estimate a difference-in-difference propensity score matching, where the treated firms are those that switch from focused to conglomerate, and the control firms are focused firms that never change their status. The results are table 10, while the details of the variables used in propensity score models are in the appendix C (C.2). Columns (1)-(4) of table 10 report the difference in difference estimation on the treated firms that start to diversify, one year before and after the diversification, with the control sample being the matched focused firms in the same period.

Each estimation is performed according to two propensity score models which estimates the propensity to diversify: the reduced model and the enhanced model. To illustrate, column (1) reports the difference-in-difference estimation according to the reduced model, while column (2) reports the difference-in-difference estimation according to the propensity score enhanced model. In columns (3) and (4), we estimate a triple difference propensity score matching where our interaction variable is "low survival", an indicator variable equal to one if the firm has a default probability above the median in the year before the change of status from focused to conglomerate. The results confirm that there is no premium associated with low survival conglomerates at the stage of conglomerate formation. In columns (3) and (4), we see that firms becoming conglomerates have the same value in the year after the switch (first row). This also holds true for high-survival conglomerates that have a similar value after (second row) and before (third row) the switch. Consistent with Proposition 1a, all firms with lower survival probabilities display lower values.

In more detail, the coefficient of "switch status×after" shows that the excess value of focused firms that become conglomerates relative to firms that remain focused does not change after the switch. In turn, the coefficient of "switch status×after×low survival" measures whether the excess value is any lower for firms that switch with a low prior survival probability relative to their focused peers that also have low survival probabilities. The coefficient is negative (-0.004) but is not statistically different from zero. The coefficient of "switch status×low survival" also indicates that the excess value for low-survival focused companies that switch is no higher than for low-survival focused companies that do not switch. The coefficient "low survival×after" shows the excess value changes after the event for low-survival firms, in general. This coefficient is, again, not statistically different from zero. Finally, the coefficient of "low survival" shows a discount of 34% for all low-survival firms (both before/after and

[13] Villalonga (2004a) finds 150 firms in a sample from 1978–1997.

switching/not switching), consistent with Proposition 1a. The coefficient of "treated" indicates the excess value gain from shifting status relative to the value of the control group of focused firms that did not switch. Companies that switched lost some value relative to the value they would have had as focused companies, but the change is not statistically significant.

In Proposition 3b, the ex-ante value of high survival firms is positive, for positive bankruptcy costs. One may argue that the lack of ex-ante premium at stage 0 in the previous results can also imply zero bankruptcy costs. However, the longitudinal test cannot test proposition 3b, as it considers a very limited portion of the population of conglomerate firms (only 150), making the sample of conglomerates very unpaired when compared to single segment firms. In order to test proposition 3b, we should observe the total sample of conglomerate population at the time of diversification, which often happens far before the firm' IPO.

In a second step, we only keep these firms in our sample to see whether, after the time passing of these firms, similar patterns that our baseline results appear.[14] We regress company excess value on conglomerate dummy, as in equation (9), before and after the correction for the survivorship bias. Table 11 reports the results. Before the correction, firms that switch from focused to conglomerate firms experience a value discount in the years following the switch. However, once controlling for their superior ability to survive over the years, the discount disappears and turns into a premium.

5 Summary and Conclusion

In this paper, we point out that the survivorship bias increases the value of survivors in proportion to their default probability. In turn, this implies that the price premium (discount) between survivors with lower and higher default probability will be smaller (larger) in the sample than in the population. Our model also shows conditions such that the higher survival ability of conglomerates results in an apparent discount relative to focused peers whenever the differential survival is not controlled for. These paradoxes are due to a known problem, namely the ex-post selection bias. Because databases do not contain current price information on the (focused) firms that disappeared in a downturn due to defaults, while they do include their (diversified) peers that survived, the ex-post relative average price does not reflect the ex-ante value of the population of firms.

We test a set of predictions in a sample of US companies from 1980 to 2014. We exploit the id-

[14] Villalonga (2004b) also shows that the discount appears already after two years.

iosyncratic difference in survival probability, both among conglomerates and between conglomerates and focused industry peers, to study the sign of the correlation between (the differences in) firms' survival probability and firms' equity values. Consistent with the survivorship bias hypothesis, this correlation is negative: the higher is the (difference in) survival probability, the higher the sample discount, after controlling for observable characteristics. Our empirical analysis shows that such patterns disappear when either there is no survivorship bias or we undo the bias by controlling for firms' survival ability,

Our analysis has implications over and beyond conglomerate value. It implies that cross-sectional comparisons aimed at assessing differences in ex-ante values should correct for differential survival. Furthermore, our investigation follows the original insight of Brown et al. (1995), but looks at the cumulative effects of many individual company defaults rather than market crashes. Using this logic, the survivorship bias in the equity premium may then appear larger than currently assessed. We leave these challenging extensions for further research.

References

Almeida, H., Philippon, T., 2007. The risk-adjusted cost of financial distress. The Journal of Finance 62, 2557–2586.

Alonso, R., Dessein, W., Matouschek, N., 2015. Organizing to adapt and compete. American Economic Journal: Microeconomics 7, 158–187.

Anderson, C.W., Garcia-Feijoo, L., 2006. Empirical evidence on capital investment, growth options, and security returns. The Journal of Finance 61, 171–194.

Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91, 444–455.

Avramov, D., Chordia, T., Jostova, G., Philipov, A., 2016. Excess value and restructurings by diversified firms. Journal of Banking & Finance 71, 1–19.

Banal-Estañol, A., Ottaviani, M., Winton, A., 2013. The flip side of financial synergies: Coinsurance versus risk contamination. The Review of Financial Studies 26, 3142–3181.

Banz, R.W., Breen, W.J., 1986. Sample-dependent results using accounting and market data: Some evidence. The Journal of Finance 41, 779.

Berger, P.G., Ofek, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37, 39–65.

Bessembinder, H., 2018. Do stocks outperform treasury bills? Journal of Financial Economics 129, 440–457.

Boot, A.W., Schmeits, A., 2000. Market discipline and incentive problems in conglomerate firms with applications to banking. Journal of Financial Intermediation 9, 240–273.

Borghesi, R., Houston, J., Naranjo, A., 2007. Value, survival, and the evolution of firm organizational structure. Financial Management 36, 5–31.

Brown, S.J., Goetzmann, W.N., Ross, S.A., 1995. Survival. The Journal of Finance 50, 853–873.

Campa, J.M., Kedia, S., 2002. Explaining the diversification discount. The Journal of Finance 57, 1731–1762.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. The Journal of Finance 63, 2899–2939.

Chava, S., Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. Review of Finance 8, 537–569.

Custodio, C., 2014. Mergers and acquisitions accounting and the diversification discount. The Journal of Finance 69, 219–240.

Duan, J.C., Sun, J., Wang, T., 2012. Multiperiod corporate default prediction: A forward intensity approach. Journal of Econometrics 170, 191–209.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fulghieri, P., Sevilir, M., 2011. Mergers, spin-offs and employee incentives. Review of Financial Studies 24, 2207–2241.

Glaser, M., Mueller, S., 2010. Is the diversification discount caused by the book value bias of de. The Journal of Banking and Finance 68.

Gomes, J., Livdan, D., 2004. Optimal diversification: Reconciling theory and evidence. The Journal of Finance 59, 507–535.

Graham, J.R., Lemmon, M., Wolf, J.G., 2002. Does corporate diversification destroy value? Journal of Finance 57, 695–720.

Hann, R.N., Ogneva, M., Ozbas, O., 2013. Corporate diversification and the cost of capital. Journal of Finance 54, 778–885.

Hennessy, C.A., Whited, T.M., 2007. How costly is external financing? evidence from a structural estimation. The Journal of Finance 62, 1705–1745.

Julliard, C., Ghosh, A., 2012. Can rare events explain the equity premium puzzle? Review of Financial Studies 25.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.

Kothari, S.P., Shanken, J., Sloan, R.G., 1995. Another look at the cross-section of expected stock returns. Journal of Finance 50, 185–224.

Kuppuswamy, V., Villalonga, B., 2016. Does diversification create value in the presence of external financing constraints? evidence from the 2007–2009 financial crisis. Management Science 62, 905–923.

Lamont, O., Polk, C., 2002. Does diversification destroy value? evidence from the industry shocks. Journal of Financial Economics 63, 51–77.

Lang, L., Stulz, R., 1994. Tobin's q, corporate diversification, and firm performance. Journal of Political Economy 102, 1248–80.

Lee, B.S., Li, M.Y.L., 2012. Diversification and risk-adjusted performance: A quantile regression approach. Journal of Banking & Finance 36, 2157–2173.

Leland, H.E., 2007. Financial synergies and the optimal scope of the firm: Implications for mergers, spinoffs, and structured finance. The Journal of Finance 62, 765–807.

Lewellen, W.G., 1971. A pure financial rationale for the conglomerate merger. Journal of Finance 26, 521–37.

Li, H., Xu, Y., 2002. Survival bias and the equity premium puzzle. The Journal of Finance 57, 1981–1995.

Luciano, E., Nicodano, G., 2014. Guarantees, leverage, and taxes. Review of Financial Studies 27, 2736–2772.

Mansi, S., Reeb, D., 2002. Corporate diversification: What gets discounted? The Journal of Finance 57, 2167–2183.

Matvos, G., Seru, A., Silva, R.C., 2018. Financial market frictions and diversification. Journal of Financial Economics 127, 21–50.

Rajan, R., Servaes, H., Zingales, L., 2000. The cost of diversity: The diversification discount and inefficient investment. The Journal of Finance 55, 35–80.

Ramey, V.A., Shapiro, M.D., 1998. Capital churning. Working paper .

Rudolph, C., Schwetzler, B., 2014. Mountain or molehill? downward biases in the conglomerate discount measure. Journal of Banking & Finance 40, 420–431.

Santioni, R., Schiantarelli, F., Strahan, P.E., 2020. Internal capital markets in times of crisis: The benefit of group affiliation. Review of Finance 24, 773–811.

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model. Journal of Business 74, 101–124.

Stein, J., 2002. Information production and capital allocation: Decentralized versus hierarchical companies. Journal of Finance 57, 1891–1921.

Villalonga, B., 2004a. Diversification discount or premium? new evidence from the business information tracking series. The Journal of Finance 59, 479–506.

Villalonga, B., 2004b. Diversification discount or premium? new evidence from the business information tracking series. The Journal of Finance 59, 479–506.

Whited, T.M., 2001. Is it inefficient investment that causes the diversification discount? The Journal of Finance 56, 1667–1691.

Yan, A., Yang, Z., Jiao, J., 2010. Conglomerate investment under various capital market conditions. Journal of Banking & Finance 34, 103–115.

Table 1: Number of companies per year

This table lists the total number of active companies, the number of active conglomerates, defaults, failures, new entries and exits of firms. One observation is at firm-year level. We define default an indicator variable equal to one if the firms defaults in a specific year. We retrieve default information from Compustat North America (delisted, bankruptcy filings and liquidations), CRSP (delisted due to a distress event), and from the UCLA-LoPucki Bankruptcy Research Database (Chapter 7 and Chapter 11). The failure events also include firms with financial default or with a D rating. The cumulative default column captures the the number of cumulative events of failure from the beginning of the sample. We define new entries as companies with end of period gross capital not bigger than 20% of the end of period net capital during the company's first year in the data set (as in Ramey and Shapiro [1998]). We define other exits as firms that exit the sample for unknown reasons, different from default, liquidation, or mergers. The sample period includes all non-financial and non-utility firms in the US, over the years 1980 - 2014.

Year	Active firms	Conglomerates	Default	Failures	Cum. Default	Mergers	New entries	Other exits
1980	2,016	1,105	23	23	23	181	0	1,462
1981	2,042	1,097	19	19	42	177	108	1,489
1982	2,085	1,055	22	22	64	199	162	1,490
1983	2,150	1,020	23	23	87	241	261	1,494
1984	2,271	992	31	31	118	248	386	1,573
1985	2,278	932	24	24	142	272	365	1,538
1986	2,319	870	24	24	166	256	478	1,538
1987	2,476	834	32	32	198	290	626	1,591
1988	2,450	769	43	43	241	239	651	1,626
1989	2,362	725	56	57	298	446	612	1,374
1990	2,364	717	52	53	351	390	668	1,419
1991	2,405	711	44	46	397	441	662	1,406
1992	2,561	739	30	33	430	524	781	1,457
1993	2,813	742	26	26	456	651	930	1,514
1994	3,087	742	48	49	505	800	1,086	1,574
1995	3,347	750	47	48	553	934	1,199	1,654
1996	3,606	752	65	67	620	1,132	1,283	1,632
1997	3,642	728	107	108	728	1,216	1,333	1,561
1998	3,313	1,124	134	141	869	1,064	1,170	1,383
1999	2,557	870	110	112	981	725	999	1,122
2000	2,327	629	86	93	1,074	551	969	1,071
2001	2,029	602	64	70	1,144	453	710	879
2002	1,859	545	30	37	1,181	428	605	758
2003	1,692	509	18	24	1,205	452	574	603
2004	1,674	506	22	25	1,230	468	644	547
2005	1,643	502	22	23	1,253	484	669	484
2006	1,627	504	33	34	1,287	468	692	415
2007	1,629	479	45	48	1,335	391	698	397
2008	1,519	465	37	39	1,374	304	549	309
2009	1,419	434	25	26	1,400	339	502	235
2010	1,407	429	27	29	1,429	373	537	164
2011	1,359	447	26	27	1,456	397	581	116
2012	1,334	438	20	22	1,478	393	554	70
2013	1,341	441	24	27	1,505	405	585	32
2014	1,386	451	20	21	1,526	406	555	2
Total	76,389	24,655	1,459	1,526	1,526	16,738	23,184	35,979

Table 2: Univariates

The table reports statistics for all variables used in the sample. Panel A reports the statistics for company value, default, and financial characteristics across company type (conglomerates vs. focused companies), and tests for univariate differences. Panel B reports the univariate statistics of the main variables used in the regressions according to 10^{th}, 25^{th}, and 50^{th} percentiles of companies' survival probability, and the statistical t-test of average differences between conglomerates and focused firms for each group. The details of the variables are in Appendix A.2. The sample consists of the intersection of the Compustat, CRSP, and the UCLA- LoPucki Bankruptcy Research Database (BRD) over the years 1980 - 2014. The test difference between conglomerates and focused companies are estimated with an OLS regression, clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Panel A	Focused		Conglomerates			
	Mean	Sd	Mean	Sd	Difference	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
Excess value	-0.006	0.649	-0.066	0.641	-0.060***	(-12.06)
Survival probability	98.09	3.322	98.26	2.944	0.165***	(6.64)
Excess PD (estimated)	-0.033	1.199	-0.271	1.357	-0.238***	(-24.53)
Excess PD (CRI)	-0.009	0.724	-0.108	0.769	-0.099***	(-7.06)
Default (Y)	0.021	0.144	0.015	0.121	-0.006***	(-5.82)
Mergers (Y)	0.228	0.420	0.199	0.400	-0.029***	(-9.08)
New entries	0.348	0.476	0.210	0.407	-0.138***	(-39.11)
CFCORR	0.962	0.211	0.416	0.577	-0.546***	(-190.32)
Leverage	0.248	0.190	0.263	0.168	0.825***	(60.26)
Size	5.269	1.675	6.094	1.953	7.365***	(80.73)
Age	13.157	11.013	20.522	13.271	0.008***	(9.50)
EBITDA	0.116	0.124	0.125	0.094	-0.006***	(-8.56)
Capex	0.077	0.095	0.071	0.072	-0.039***	(-19.18)
Sales Growth (SG)	0.152	0.273	0.113	0.246	0.005***	(29.70)
Dividend ratio	0.009	0.021	0.014	0.020	0.015***	(10.49)
Obs.	51,734		24,655			

Table 2: Univariates - continued

Panel B: Survival skills quintiles	10%			25%			50%		
	Mean	Sd	Diff	Mean	Sd	Diff.	Mean	Sd	Diff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Excess value	-0.279	0.600	0.056**	-0.270	0.592	0.010	-0.155	0.592	-0.052***
Size	5.204	1.699	0.573***	5.228	1.764	0.711***	5.245	1.689	0.619***
Age	14.662	10.572	3.594***	15.149	11.132	4.483***	15.761	11.573	5.997***
EBITDA	0.037	0.125	0.019***	0.078	0.106	0.014***	0.110	0.106	0.007***
Capex	0.059	0.081	-0.005**	0.068	0.087	-0.006***	0.073	0.088	-0.006***
Sales Growth (SG)	0.051	0.275	-0.002	0.107	0.268	-0.012**	0.135	0.260	-0.028***
Dividend ratio	0.004	0.013	0.001***	0.006	0.017	0.003***	0.009	0.020	0.003***
Leverage	0.437	0.190	-0.008	0.350	0.175	-0.001	0.267	0.162	0.006*
Obs	8,523			12,683			20,194		

Table 3: Excess value and excess default probability: multivariate regression

The table reports the results of the estimation of the following equation:

$$y_{i,t} = \alpha + \beta\, Conglomerate_{it} + \Gamma X_{i,t-1} + \varepsilon_t \,,$$

where the dependent variables are the excess value and the excess default probability, over the years 1980 - 2014, of conglomerates and focused firms. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The excess default probability is computed as the natural logarithm of the ratio between a company's PD (one minus its survival probability) and its imputed PD at the end of the year. The survival probability is the average of the values of the segments' survival, the latter being computed by multiplying the segments' sales to the median survival-to-sales multiplier of the single-segment companies in the same industry of the segment unit, attributed by using the narrower SIC code. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model controls for a vector of company characteristics (listed in the table), including year and industry (Fama-French 17) fixed effects. In columns (5)-(6), we run the estimation after excluding firms involved in any corporate event that affects the sample composition: new entries, exits, failures, mergers. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	All Sample				Reduced Sample	
	Excess value		Excess PD		Excess value	Excess PD
	(1)	(2)	(3)	(4)	(3)	(4)
Conglomerate	-0.122***	-0.094***	-0.072***	-0.065***	-0.100***	-0.140***
	(-10.618)	(-7.966)	(-4.220)	(-3.750)	(-4.160)	(-3.785)
Age		-0.088***		-0.021**	-0.053***	-0.087***
		(-12.495)		(-2.029)	(-3.231)	(-3.595)
Leverage		0.041	2.195***	2.188***	0.031	2.361***
		(1.55)	(55.21)	(55.14)	(0.50)	(24.29)
Assets	0.079***	0.089***	-0.225***	-0.222***	0.089***	-0.248***
	(21.74)	(22.61)	(-36.029)	(-35.141)	(11.91)	(-19.742)
CAPEX	0.277***	0.305***	-0.644***	-0.636***	0.616***	-0.991***
	(5.91)	(6.47)	(-9.731)	(-9.578)	(5.66)	(-6.352)
Sales growth	0.358***	0.305***	-0.227***	-0.240***	0.315***	-0.308***
	(28.06)	(23.58)	(-12.331)	(-12.842)	(10.40)	(-6.904)
Dividends	1.428***	1.537***	-3.501***	-3.487***	2.527***	-4.517***
	(5.56)	(5.81)	(-10.086)	(-10.081)	(4.91)	(-6.072)
EBITDA			-1.568***	-1.569***	-0.075	-1.542***
			(-19.321)	(-19.346)	(-0.641)	(-7.757)
NITA			-1.295***	-1.288***	0.000	-1.500***
			(-21.324)	(-21.211)	0.00	(-9.922)
CACL			-0.125***	-0.126***	0.000	-0.127***
			(-29.144)	(-29.093)	0.00	(-13.941)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.080	0.090	0.371	0.371	0.128	0.412
N	75,393	75,393	75,393	75,393	15,252	15,252

Table 4: Conglomerate Discount and Excess Default by Survival Probability

The table reports of the following equation:

$$y_{i,t} = \alpha + \beta\, Conglomerate_{it} + \Gamma X_{i,t-1} + \varepsilon_t$$

where the dependent variables is the excess value over the years 1980 - 2014, computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC including at least five single-segment companies. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model is performed on four subsamples, split according to the 10^{th}, 25^{th}, 50^{th}, and 100^{th} percentiles of companies' survival probability, as computed in Campbell et al. (2008). The model controls for a vector of company characteristics (listed in the table), including year and industry fixed effects. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Dep Var: ExcessValue				Dep Var: Excess Default PD			
Survival Probability Q:	p(10) (1)	p(25) (2)	p(50) (3)	p(100) (4)	p(10) (5)	p(25) (6)	p(50) (7)	p(100) (8)
Conglomerate	-0.013	-0.027*	-0.053***	-0.109***	0.286***	0.102***	-0.078***	-0.290***
	(-0.711)	(-1.648)	(-3.794)	(-6.789)	(10.70)	(4.95)	(-4.346)	(-13.978)
Age	-0.033***	-0.028***	-0.061***	-0.119***	0.010	0.006	0.016	-0.004
	(-2.583)	(-2.959)	(-7.388)	(-13.061)	(0.56)	(0.49)	(1.57)	(-0.297)
Assets	0.109***	0.067***	0.040***	0.061***	-0.081***	-0.063***	-0.054***	-0.219***
	(17.38)	(13.13)	(8.62)	(11.08)	(-9.513)	(-9.671)	(-9.015)	(-29.757)
EBITDA	-0.762***	-1.206***	-1.136***	-0.381***	-0.540***	-0.428***	-0.449***	-0.619***
	(-11.353)	(-17.143)	(-18.402)	(-5.886)	(-5.081)	(-3.783)	(-4.294)	(-5.679)
CAPEX	0.165*	0.251***	0.084	0.170***	0.028	-0.229**	-0.171**	-0.186**
	(1.82)	(3.46)	(1.39)	(2.70)	(0.22)	(-2.503)	(-2.175)	(-2.223)
Sales growth	0.024	0.075***	0.156***	0.358***	0.021	0.008	0.044*	-0.119***
	(0.98)	(3.64)	(8.54)	(19.94)	(0.63)	(0.29)	(1.84)	(-5.016)
Dividends	-0.604	0.123	0.549*	1.517***	-0.139	-0.687	-0.584	-1.170***
	(-1.238)	(0.300)	(1.658)	(5.652)	(-0.189)	(-1.429)	(-1.589)	(-3.256)
Chi2				12.51				26.50
pvalue				0.000				0.000
Industry FE		Yes	Yes	Yes		Yes	Yes	Yes
Year FE		Yes	Yes	Yes		Yes	Yes	Yes
R-squared	0.296	0.252	0.184	0.081	0.181	0.112	0.085	0.202
N	8,510	12,785	21,587	43,481	8,510	12,785	21,587	43,481

Table 5: Instrumental Variables (IV) Approach

The table reports the results of the estimation of our baseline model after controlling through IV for excess default probability. We estimate 9 after correcting our sample according to equation 11, therefore using as dependent variable the residuals from regressing the excess value on the two, three, and four periods lagged excess default probability, respectively. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The excess default probability is computed as the natural logarithm of the ratio between a company's PD (one minus its survival probability) and its imputed PD at the end of the year. The survival probability is the average of the values of the segments' survival, the latter being computed by multiplying the segments' sales to the median survival-to-sales multiplier of the single-segment companies in the same industry of the segment unit, attributed by using the narrower SIC code. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model controls for a vector of company characteristics (listed in the table), including year and industry (Fama-French 17) fixed effects. The model controls for a vector of company characteristics used throughout, including year and firm fixed effects. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels

	All sample						Reduced sample		
Instrument:	Two lags		Three lags		Four lags		Two lags	Three lags	Four lags
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Conglomerate	0.006*	0.007**	0.004	0.005*	0.003	0.004	0.003	0.003	0.001
	(1.89)	(2.35)	(1.39)	(1.82)	(0.91)	(1.33)	(1.06)	(0.83)	(0.41)
Age		-0.001		-0.002	0.000	-0.002	-0.004*	-0.006**	-0.006**
		(-0.338)		(-0.786)	0.00	(-0.784)	(-1.891)	(-2.374)	(-2.317)
Leverage		-0.278***		-0.221***	0.000	-0.180***	-0.272***	-0.216***	-0.178***
		(-40.015)		(-31.875)	0.00	(-26.307)	(-37.463)	(-29.722)	(-24.511)
Assets	0.029***	0.034***	0.027***	0.031***	0.026***	0.028***	0.032***	0.030***	0.027***
	(25.73)	(30.05)	(25.12)	(28.58)	(24.60)	(27.36)	(27.97)	(26.42)	(25.16)
CAPEX	0.126***	0.164***	0.047***	0.075***	-0.014	0.008	0.161***	0.078***	0.005
	(8.06)	(11.75)	(3.01)	(5.28)	(-0.901)	(0.54)	(10.70)	(4.96)	(0.28)
Sales growth	0.022***	0.025***	0.007**	0.010***	-0.001	0.002	0.026***	0.010**	0.001
	(5.65)	(6.69)	(1.97)	(2.74)	(-0.190)	(0.54)	(6.17)	(2.46)	(0.13)
Dividends	1.210***	0.941***	1.114***	0.899***	1.006***	0.832***	0.802***	0.783***	0.717***
	(14.00)	(12.24)	(13.35)	(11.90)	(12.01)	(10.88)	(9.60)	(9.41)	(8.62)
EBITDA			0.115***	0.049***	0.097***	0.041***	0.078***	0.040***	0.041***
			(9.15)	(4.12)	(7.50)	(3.29)	(6.12)	(3.10)	(3.05)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.168	0.242	0.162	0.219	0.162	0.206	0.216	0.193	0.181
N	50,989	50,989	43,803	43,803	37,887	37,887	38,309	32,462	27,745

Table 6: Instrumental Variables Approach: Quantile Regression

The table reports the results of the estimation of our baseline model after correcting for the firms superior ability to survive. We estimate 9 on four subsamples, split according to the 10^{th}, 25^{th}, 50^{th}, and 100^{th} percentiles of companies' survival probability (the latter computed as in Campbell et al. (2008)), after correcting our sample according to equation 11, therefore using as dependent variable the residuals from regressing the excess value on the two-periods lagged. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The excess default probability is computed as the natural logarithm of the ratio between a company's PD (one minus its survival probability) and its imputed PD at the end of the year. The survival probability is the average of the values of the segments' survival, the latter being computed by multiplying the segments' sales to the median survival-to-sales multiplier of the single-segment companies in the same industry of the segment unit, attributed by using the narrower SIC code. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model controls for a vector of company characteristics (listed in the table), including year and industry (Fama-French 17) fixed effects. The model controls for a vector of company characteristics used throughout, including year and firm fixed effects. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels

Panel B: Quantile Regression				
	Survival Probability Distribution			
	p(10)	p(25)	p(50)	p(100)
	(1)	(2)	(3)	(4)
Conglomerate	-0.021***	-0.004	-0.001	0.031***
	(-3.646)	(-0.846)	(-0.218)	(7.88)
Age	-0.022***	-0.011***	-0.012***	-0.002
	(-5.887)	(-3.497)	(-4.414)	(-0.725)
Assets	0.014***	0.015***	0.020***	0.036***
	(7.50)	(10.29)	(14.37)	(25.69)
EBITDA	-0.179***	-0.196***	-0.094***	0.110***
	(-8.359)	(-8.718)	(-4.880)	(5.79)
CAPEX	0.135***	0.195***	0.162***	0.066***
	(4.65)	(8.87)	(8.10)	(3.27)
Sales growth	0.017*	0.023***	0.011*	0.003
	(1.91)	(3.27)	(1.69)	(0.49)
Dividends	0.017	0.420***	0.589***	0.672***
	(0.091)	(3.401)	(6.024)	(7.080)
Chi2				7.66
pvalue				0.007
Industry FE		Yes	Yes	Yes
Year FE		Yes	Yes	Yes
R-squared	0.078	0.074	0.079	0.199
N	5,904	8,500	13,217	23,368

Table 7: Excess value of conglomerate firms: Fixed effects

The table reports the results of the estimation of equation 9, where the dependent variable is firms' the excess value, over the years 1980 - 2014, of conglomerates and focused firms. In columns (1)-(2), the model is performed on on the raw data, while columns (3)-(4) report the estimations when using as dependent variable the residual of a regression of the excess value on the lagged values of the excess default probability, as in equation 11. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies. The excess default probability is computed as the natural logarithm of the ratio between a company's PD (one minus its survival probability) and its imputed PD at the end of the year. The survival probability is the average of the values of the segments' survival, the latter being computed by multiplying the segments' sales to the median survival-to-sales multiplier of the single-segment companies in the same industry of the segment unit, attributed by using the narrower SIC code. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model controls for a vector of company characteristics used throughout, including year and firm fixed effects. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels

	Before Correction		After correction	
	(1)	(2)	(3)	(4)
Conglomerate	-0.105***	-0.088***	0.001	0.004
	(-7.920)	(-6.675)	(0.170)	(10.959)
Age		-0.263***		-0.053***
		(-18.929)		(-7.727)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R-squared	0.594	0.606	0.528	0.542
N	76,389	76,389	51,712	51,712

Table 8: Excess value and measures of diversification

The table reports the estimates of equation 9 with different proxies of the independent variable, where the dependent variable is the excess value, before (columns (1)-(3), and after (columns (4)-(6)) the correction for the firms survivorship bias, as in equation 11. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We use different proxies for the variable diversification: the number of segments, the coinsurance across segments (one minus the segment cash-flow correlation), and the number of three digits SIC code industries in which the firm is operating. The model controls for the vector of company characteristics used throughout, including year and industry fixed effects. The sample include all non-financial, and non-utility firms, over the years 1980-2014. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. Var: Excess value	Before correction for PD			After correction for PD		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of segments	-0.102***			0.005		
	(-9.045)			(1.47)		
Coinsurance		-0.029***			0.005	
		(-2.664)			(1.50)	
Number of industies			-0.038***			0.005***
			(-6.008)			(2.59)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.091	0.086	0.088	0.169	0.169	0.169
N	75,393	75,393	75,393	50,989	50,989	50,989

Table 9: Excess value by leverage groups

The table reports the estimation of equation 9 by leverage groups, where the dependent variables is the excess value, over the years 1980 - 2014, regressed within leverage quintiles. In columns (1)-(3), we perform the estimation in the raw data, while in columns (4)-(6), we performs the estimation after correcting for the survovorship bias as in equation 11. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The variable "conglomerate" is an indicator variable equal to one if the company is multi-segments. The model is performed on four sub-samples split according to the 10^{th}, 25^{th}, and 50^{th} percentiles of the company leverage. The model controls for a vector of company characteristics used throughout, including year and industry fixed effects. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Before Correction			After correction		
Leverage Distribution:	p(10) (1)	p(25) (2)	p(50) (3)	p(10) (4)	p(25) (5)	p(50) (6)
Conglomerate	-0.129***	-0.108***	-0.065***	0.018**	0.018***	0.000
	(-3.519)	(-6.594)	(-4.736)	(2.25)	(3.92)	(0.01)
Age	-0.162***	-0.086***	-0.068***	-0.022***	0.003	0.006**
	(-7.699)	(-8.215)	(-8.480)	(-3.639)	(0.87)	(2.21)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.112	0.091	0.106	0.219	0.206	0.179
N	7,594	30,289	37,510	4,316	20,659	26,014

Table 10: Excess value at the ex-ante stage of conglomerate formation

The table reports the results of the estimation of the following equation:

$$y_{i,t} = \alpha + \beta\, treated_i + \beta_1\, treated_i \times after_t + \Gamma X_{i,t-1} + \varepsilon_t$$

where the dependent variable is the excess value of treated and control groups. The treated group is composed by firms that change their status from focused to conglomerate firms (multisegment firms), while focused firms compose the control group. We estimate the difference-in-difference regression as in Villalonga (2004a)) over a window of one year before/after the change. The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The model controls for a vector of company characteristics used throughout, including year and firm fixed effects. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels

Dep. var.: Excess Value	(1)	(2)	(3)	(4)
Switch status×after	0.012	0.001	0.010	-0.002
	(0.437)	(0.025)	(0.320)	(-0.048)
Switch status×after×low survival			-0.004	0.016
			(-0.060)	(0.234)
Switch status×low survival			-0.015	-0.008
			(-0.217)	(-0.121)
Low survival×after			0.010	-0.011
			(0.350)	(-0.374)
Treated	-0.048	-0.041	-0.020	-0.004
	(-1.462)	(-1.281)	(-0.343)	(-0.068)
Low survival			-0.340***	-0.345***
			(-10.650)	(-11.178)
Propensity score model	Reduced	Enhanced	Reduced	Enhanced
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adj-R2	0.094	0.092	0.139	0.141
N	30,554	30,516	30,441	30,441

Table 11: Ex-post value correction of switching firms

The table reports the results of the estimation of our baseline model after keeping in the sample firms that switch from single-segment to conglomerates, as in table 10, and their control sample, before and after the correction for their excess default probability. The dependent variable is the excess value of treated and control groups over the entire sample. The treated group is composed by firms that change their status from focused to conglomerate firms (multisegment firms), while focused firms compose the control group. In columns (1)-(3), we estimate 9 before any correction of the sample for the firms default probability, therefore using as dependent variable the residuals from regressing the excess value on the two-periods lagged excess default probability.The excess value is computed as the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC code, for industries including at least five single-segment companies in the year of the analysis. The excess default probability is computed as the natural logarithm of the ratio between a company's PD (one minus its survival probability) and its imputed PD at the end of the year. The survival probability is the average of the values of the segments' survival, the latter being computed by multiplying the segments' sales to the median survival-to-sales multiplier of the single-segment companies in the same industry of the segment unit, attributed by using the narrower SIC code. The model controls for a vector of company characteristics used throughout, including year and firm fixed effects. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels

Dep. Var: Excess value	Before correction for PD			After correction for PD		
	(1)	(2)	(3)	(4)	(5)	(6)
Conglomerate	-0.116***	-0.065***	-0.066***	0.005	0.011*	0.010
	(-5.532)	(-3.023)	(-3.123)	(0.81)	(1.85)	(1.33)
Age		-0.104***	-0.262***		-0.013***	-0.073***
		(-9.572)	(-13.680)		(-3.804)	(-6.866)
Leverage		-0.035	0.029		-0.301***	-0.195***
		(-0.928)	(0.74)		(-28.641)	(-13.347)
Assets	0.085***	0.096***	0.117***	0.032***	0.038***	0.072***
	(15.43)	(16.08)	(11.57)	(18.37)	(22.60)	(17.85)
CAPEX	0.425***	0.443***	0.523***	0.119***	0.154***	0.179***
	(6.38)	(6.66)	(11.78)	(5.29)	(7.72)	(9.20)
Sales growth	0.361***	0.303***	0.072***	0.027***	0.023***	-0.004
	(19.92)	(16.48)	(5.10)	(4.36)	(3.94)	(-0.781)
Dividends	1.510***	1.654***	1.210***	1.111***	0.906***	0.792***
	(4.02)	(4.32)	(4.85)	(8.85)	(7.95)	(6.33)
EBITDA	0.083	0.093	0.340***	0.183***	0.101***	-0.080***
	(1.19)	(1.33)	(5.84)	(9.40)	(5.55)	(-4.243)
Firm FE	No	No	Yes	No	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.104	0.118	0.599	0.153	0.232	0.523
N	34,483	34,483	34,483	25,997	25,997	25,997

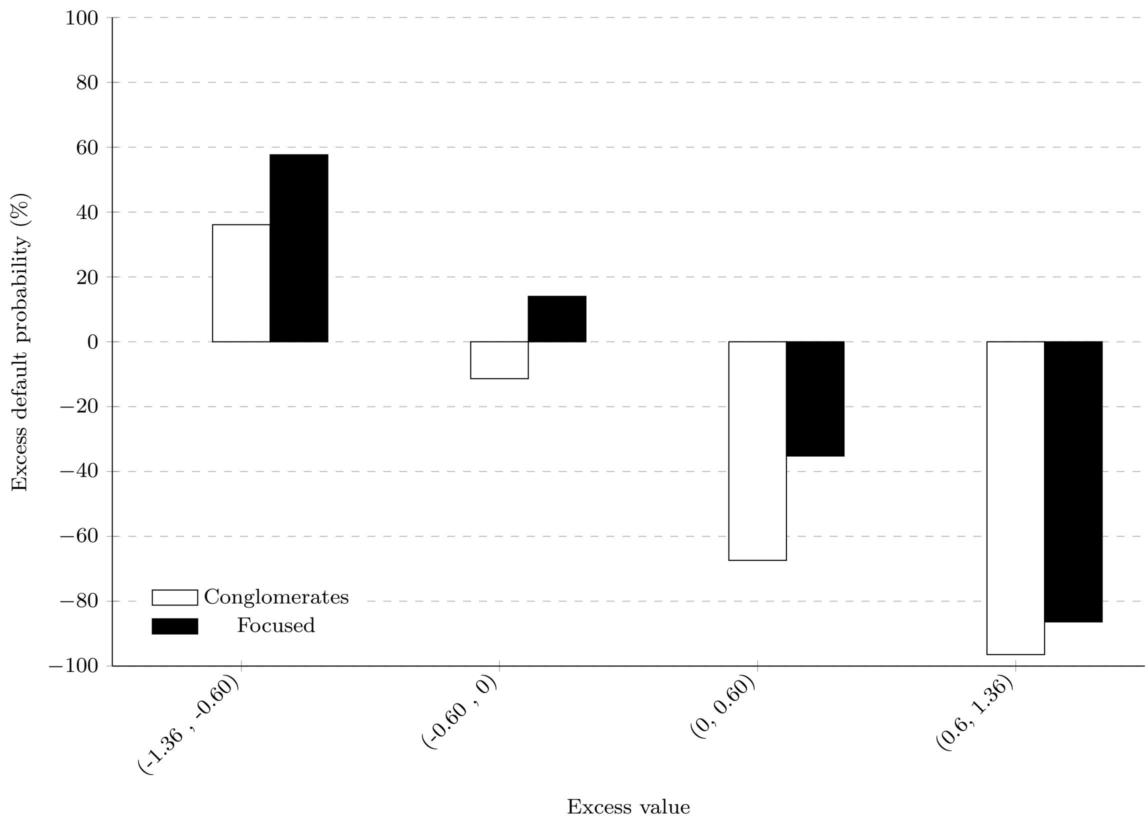

Figure 1: Excess default probability by excess values categories

This figure reports the excess probability of default of conglomerates and focused companies for different intervals of the excess value. The excess value is the natural logarithm of the ratio between a company's market value and its imputed value at the end of the year. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. We implement the industry matching using the narrower SIC including at least five single-segment companies. For each interval of the computed excess value, we report the value of the excess probability of default, computed as the natural logarithm of the ratio between a company PD and its imputed PD at the end of the year. The company PD is computed following Cambpell et al (2008), as reported in Table 2. The imputed PD is the average of the values of the segments' PD, the latter being computed by multiplying the segments' sales to the median PD-to-sales multiplier of the single-segment companies in the same industry as the segment unit. We retrieve information on company bankruptcy from Compustat North America database, CRSP, and UCLA- LoPucki Bankruptcy Research Database (BRD). The sample period goes from 1980 to 2014.

A Proofs and Extensions

A.1 Lemma and Proofs of All Propositions

Lemma: State Space and Borrowing Costs:

Assume $D_B p_B^{-1} \leq X_B < (D_A + D_B)[p_A + p_B(1 - p_A)]^{-1}$; and

$X_A \geq D_A p_A^{-1} + D_B[p_B + p_A(1 - p_B)]^{-1}$.

Then:

a. the state space is $\{HH, LL, HL, LH\}$, as defined in the main text;

b. the following ranking of borrowing costs holds across company types:

$$R_C < R_A + R_B \tag{A.1}$$

Proof of the Lemma:

a. In state $\{H\}$, it must be the case that cash flow, X_i, exceeds the total debt repayment in each unit. For unit B, this requires that

$$X_B \geq D_B p_B^{-1} \tag{A.2}$$

In state $\{LH\}$, unit B is unable to rescue unit A. Since conglomerate lenders require a lower interest rate than focused lenders (by the ranking in Part(b)), the condition simplifies to $X_B < R_C$, that is:

$$X_B < [D_A + D_B - p_B(1 - p_A)X_B]p_A^{-1}$$
$$p_A X_B < D_A + D_B - p_B(1 - p_A)X_B$$
$$[p_A + p_B(1 - p_A)]X_B < D_A + D_B$$

which implies

$$X_B < (D_A + D_B)[p_A + p_B(1 - p_A)]^{-1}\,. \tag{A.3}$$

As for unit A, its profit in state $\{H\}$ must also exceed the combined service of debt for the two units, i.e. $X_A \geq max(R_C, R_A + R_{B\in G})$, that is:

$$X_A \geq D_A p_A^{-1} + D_B[p_B + p_A(1 - p_B)]^{-1}\,. \tag{A.4}$$

b. We need to prove that $R_C < R_A + R_B$, that is:

$$[D_A + D_B - p_B(1-p_A)X_B]p_A^{-1} < D_A p_A^{-1} + D_B p_B^{-1}$$
$$[D_B - p_B(1-p_A)X_B]p_A^{-1} < D_B p_B^{-1}$$
$$[D_B - p_B(1-p_A)X_B]p_B < D_B p_A$$
$$p_B^2(1-p_A)X_B > D_B p_B - D_B p_A$$
$$X_B > D_B(p_B - p_A)[p_B^2(1-p_A)]^{-1} \equiv X_B^* \,.$$

This inequality always holds because X_B, by (Equation (A.2)), exceeds $D_B p_B^{-1}$ which in turn exceeds X_B^*:

$$D_B(p_B - p_A)[p_B^2(1-p_A)]^{-1} < D_B p_B^{-1}$$
$$(p_B - p_A)[p_B(1-p_A)]^{-1} < 1$$
$$p_B - p_A < p_B(1-p_A)$$
$$p_A > p_A p_B \,,$$

since $p_B < 1$ holds by assumption.

Proposition 1: *ex-ante* company value

To prove Part (a), consider that value coincides with expected profit after the risk-adjusted service of debt, due to the zero risk-free rate assumption. In the case of two focused companies expected profit is equal to:

$$\begin{aligned} V_F &= p_A(X_A + K_A - R_A) + p_B(X_B + K_B - R_B) = \\ &= p_A X_A + p_A K_A - D_A + p_B X_B + p_B K_B - D_B = \\ &= \pi_A + \pi_B + p_A K_A + p_B K_B \,, \end{aligned} \tag{A.5}$$

which proves Equation (4) in the paper since $p_A^{Sur} = p_A$ and $p_B^{Sur} = p_B$. In turn, conglomerate expected profit is equal to

$$\begin{aligned} V_C &= p_A p_B(X_A + K_A + X_B + K_B - R_C) + p_A(1 - p_B)(X_A + K_A + K_B - R_C) = \\ &= p_A(X_A + K_A + K_B - R_C) + p_A p_B X_B = \\ &= p_A X_A + p_A(K_A + K_B) + p_B X_B - D_A - D_B = \\ &= \pi_A + \pi_B + p_A(K_A + K_B)\,, \end{aligned} \tag{A.6}$$

where $p_C^{Sur} = p_A$.

Proposition 2: the expected value of survivors

In order to determine the average market price of a surviving company, MV_i, at the interim stage, we need to know which companies belong to the sample in each state. Let us start with focused units. The probability of state $\{H\}$, when a focused company is alive, is equal to one because in other state it would have gone bankrupt. It follows that the average stock price of a focused company, when it is alive, is equal to the cash flow realizations net of the debt repayment state $\{H\}$; that is:

$$MV_i = X_i + K_i - R_i = \pi_i(p_i^{Sur})^{-1} + K_i\,. \tag{A.7}$$

In turn, the combined market value of two focused companies, when both are alive, is equal to

$$MV_F = \pi_A(p_A^{Sur})^{-1} + K_A + \pi_B(p_B^{Sur})^{-1} + K_B\,. \tag{A.8}$$

We similarly determine the average value of a surviving conglomerate. C survives in both state HH and in state HL. Therefore, the probability of state $\{HH\}$, when a conglomerate is alive, is lower than one:

$Pr(HH)/[Pr(HH) + Pr(HL)] = p_A p_B[p_A p_B + (1 - p_B)p_A]^{-1}$, which simplifies to p_B.

In turn, the probability of state $\{HL\}$, when the conglomerate is alive, (i.e., $Pr(HL)/[Pr(HH) + Pr(HL)]$), is equal to $(1 - p_B)$. Thus, the average market value of a surviving conglomerate will be a weighted average of the profits in those two states:

$$\begin{aligned} MV_C &= p_B(X_A + K_A + X_B + K_B - R_C) + (1 - p_B)(X_A + K_A + K_B - R_C) = \\ &= (\pi_A + \pi_B)(p_C^{Sur})^{-1} + K_A + K_B\,. \end{aligned} \tag{A.9}$$

This expression captures another way of thinking about the survivorship bias for conglomerates: they survive also through adverse states of the world thanks to coinsurance. Appropriately combining Equations (4) and (5) in the paper proves the proposition.

Proposition 3: the Survival Discount and the Contagion Premium

To prove part (a), we determine the difference in the expected value of survivors of different types. Subtracting (A.8) from (A.9) delivers the conglomerate excess value relative to focused companies:

$$\begin{aligned} MV_C - MV_F &= (\pi_A + \pi_B)(p_C^{Sur})^{-1} + K_A + K_B + \\ &- \pi_A(p_A^{Sur})^{-1} - K_A - \pi_B(p_B^{Sur})^{-1} - K_B = \\ &= \pi_A[(p_C^{Sur})^{-1} - (p_A^{Sur})^{-1}] + \pi_B[(p_C^{Sur})^{-1} - (p_B^{Sur})^{-1}] = \\ &= (p_B X_B - D_B)[p_A^{-1} - p_B^{-1}]\,, \end{aligned} \tag{A.10}$$

since $p_C^{Sur} = p_A^{Sur} = p_A$. The term in the first parenthesis is positive by assumption, hence the sign of the excess value is negative if and only if $p_A > p_B$. This condition ensures that coinsurance more than offsets contagion in conglomerates, leading to a higher probability of survival in conglomerates than in a combination of focused units.

To prove part (b), we just need to appropriately combine Equations (4), and (5) in the paper, as follows:

$$V_C - V_F = (p_C^{Sur} - p_B^{Sur})K_B, \tag{A.11}$$

This shows that the *ex-ante* expected excess value of conglomerates is an increasing function of their relative survival ability, if bankruptcy costs, K_B, are positive. On the contrary, the expected excess value of surviving conglomerates is a decreasing function of their relative survival ability, irrespective of bankruptcy costs (see Equations (A.10)). Therefore, the larger the *ex-ante* premium of a company type due to its excess survival, the larger its discount due to the survivorship bias will be.

A.2 Model with Mutual Supports

This section adds to the model in Section 2 the possibility that unit A rescues unit B. Each unit operating profit in $t = 1$ can therefore be medium, high or low. It will be medium $\{M\}$, and equal to $X_i^M > 0$, with probability $p_i^M \in (0, 1)$, it will be high $\{H\}$, and equal to $X_i^H > X_i^M$, with probability $p_i^H \in (0, 1)$, and it will be low and equal to zero with probability $p_i^L = (1 - p_i^M - p_i^H)$. Accordingly, we define nine states of the world, $\{LL, LM, ML, LH, HL, MM, MH, HM, HH\}$.

The key assumption of the general model is that the profit of each unit, in state $\{H\}$, exceeds the combined debt repayment of the two units, while, in state $\{M\}$, it is sufficient to honor its own debt obligations but not the combined service of debt. Consequently, not only unit A can rescue unit B in state $\{HL\}$ but also unit B can save unit A from bankruptcy in state $\{LH\}$, provided that they do not operate as independent entities. Setting $p_A^M = 0$, $p_A^H = p_A$, $p_B^M = p_B$, $p_B^H = 0$, $X_A^H = X_A$, $X_B^M = X_B$ leads to the original model where only unit A can rescue unit B in state $\{HL\}$.

Let us now consider, for each company type, survival probability, cost of debt and conditions on cash flows within this general setup. focused companies survive in states $\{M\}$ and $\{H\}$ with probability $p_i^{Sur} = (p_i^M + p_i^H)$ and default in state $\{L\}$. A conglomerate defaults in states $\{LL\}$, $\{LM\}$ and $\{ML\}$ when both units do not realize any profit, when unit A drags profitable unit B into bankruptcy and when unit B drags solvent unit A into bankruptcy, respectively. However, conglomerates survive when either their segments are both profitable, states $\{MM\}$, $\{MH\}$, $\{HM\}$ and $\{HH\}$, or one of their units can save the other from insolvency, states $\{LH\}$ and $\{HL\}$. Conglomerate survival probability is, therefore, equal to $p_C^{Sur} = (p_A^H + p_B^H - p_A^H p_B^H + p_A^M p_B^M)$.

Within this framework, the credit spread charged by the lenders, satisfying their zero expected profit condition, is equal to

$$R_i = D_i(p_i^M + p_i^H)^{-1} = D_i(p_i^{Sur})^{-1} \tag{A.12}$$

for a focused,

$$\begin{aligned} R_C &= (D_A + D_B - p_A^M p_B^L X_A^M - p_A^L p_B^M X_B^M)(p_A^H + p_B^H - p_A^H p_B^H + p_A^M p_B^M)^{-1} \\ &= (D_A + D_B - p_A^M p_B^L X_A^M - p_A^L p_B^M X_B^M)(p_C^{Sur})^{-1} \end{aligned} \tag{A.13}$$

for a conglomerate. As before, we can show that the following inequality holds:

$$R_C < R_A + R_B\,. \tag{A.14}$$

Let us define $\pi_A = X_A^M p_A^M + X_A^H p_A^H - D_A$ and $\pi_B = X_B^M p_B^M + X_B^H p_B^H - D_B$ as the expected current profit after the service of debt for unit A and B, respectively. Therefore, it can be shown that the value definitions (Equations (4)-(5) in the paper), stock price definitions (Equations (A.7)-(A.9)), and Propositions 1, 2, and 3 hold for the general model as well, once the reader takes into account the new definitions of both π_i and the survival probability of each company type.

This extension confirms the main results of the restricted model. Provided that contagion is less likely than coinsurance, the stock price differential between diversified and focused companies may grow even larger, since all units have the ability to rescue the other from bankruptcy.

B Appendix: Variables

This section reports the details of the variables construction, the complete distribution and the correlation matrix.

B.1 Construction of Variables

B.1.1 Dependent Variables

CONGLOMERATE is an indicator variable that is equal to one if the company engages in industry diversification.

EXCESS VALUE is computed as the natural logarithm of the ratio between a company's market value and its imputed value. The imputed value is the average of the market values of the segment units of the conglomerates, the latter computed by multiplying the segments sales to the median market-to-sales multiplier of the focused companies in the same industry of the segment unit. The industry matching is done by using the narrower SIC including at least five single-segment companies.

EXCESS DEFAULT PROBABILITY is computed as the natural logarithm of the ratio between a company's probability of default (PD) and its imputed PD at the end of the year. The PD is computed following Campbell et al. (2008). The imputed PD is the average of the values of the segments' PD, the latter being computed by multiplying the segments' sales to the median PD-to-sales multiplier of the single-segment companies in the same industry as the segment unit. The industry matching uses the narrower SIC including at least five single-segment companies. For robustness tests, default probabilities are retrieved from the Credit Research Initiative (CRI) of the University of Singapore (RMI-NUS). The CRI probabilities are built on the forward intensity model developed by Duan et al. (2012).

B.1.2 Independent Variables - Multivariate Regressions

AFTER is an indicator variable equal to one for the year following the switch of a firm from focused to diversified.CALC is the ratio of company Current assets (ca) to company Current liabilities (cl).

CAPEX is the ratio of company Capital Expenditure to company Total Assets.

CFCORR is the cross-segment cash flow correlation across segment units ($CFCORR$). This indicator may capture conglomerate diversification better than the number of conglomerate segments. Following (Hann et al., 2013), we first compute the average of the $EBITDA/assets$ ratio for all focused companies for each quarter-year. Second, we compute the industry cash flows as the residuals from a regression of the average industry cash flow of focused firms using the average cash flow of the market and the Fama and French (1993) factors for each year and industry. Next, we estimate pairwise industry correlations using the previous five-year industry cash flows for each year in the sample, and we impute the industry pairwise correlation according to the segment units and the segments' SIC codes. The cross-segment cash flow correlation for firm i in year t with n number of segments is computed as follows:

$$CFCorr_{it(n)} = \sum_{p=1}^{N} \sum_{q=1}^{N} w_{ip(j)} w_{iq(k)} \times Corr_{jk}[t-10, t-1](j,k) \tag{B.1}$$

where $w_{ip(j)}$ are the weights (sales of the segment over total firm sales) of segment p of firm i operating in industry j, and $Corr\left([t-10, t-1](j,k)\right)$ is the correlation of industry cash flows between industries j and k over the five-year period before year t. A high correlation coefficient between segment cash flows is a proxy for lower coinsurance across divisions with focused firms, at the maximum level having a correlation equal to one and zero coinsurance.

DIVIDENDS is the ratio of Dividends to Total Assets.

EBITDA is the ratio of company Earnings before Extraordinary Items to company Total Assets.

HIGH SURVIVAL is an indicator variable equal to one when the firm has a survival probability (1-PD) higher than the sample median of the year.

LEVERAGE is the ratio between total debt (dltt+dlc) and company total assets.

MB (market-to-book) is the ratio between the market value of company equity (computed by multiplying yearly closing price by the number of outstanding shares) and the book value of the equity (seq).

NITA is the ratio between company Net Income and company Total Assets.

SALES GROWTH is the yearly growth of the ratio of Sales and company Total Assets.

SIZE is the natural logarithm of company total assets.

B.1.3 Independent Variables - Survival Analysis

ADJSIZE is the logarithm of the total company assets adjusted by 10% of the difference between the market equity and the book equity of the company $[TA + 0.1(ME - BE)]$.

CASHMTA is the ration between company Cash and Short Term Investments and the sum of company Market Equity and the company Total Liabilities.

EBTA is the ratio between company Market Equity and the company Total Liabilities.

EXRET is the difference between the log gross company return in CRSP (ret), and the log gross return on the S& P Index.

MELT is the ratio between the Market Equity of the company and company Total Liabilities.

REAT is the ratio between company retained earnings and the total assets.

SIGMA is volatility of a company stock returns, computed as the annualized standard deviation of daily stock returns, averaged over 3 months, as follows:

$$SIGMA_{i,t-1,t-3} = \left(\frac{252 \times \sum_{t-1,t-2,t-3} r^2}{n-1} \right).$$

NIMTA is the ratio between company Net Income (ni in compustat) and the sum of company Market Equity to Total Liabilities (net income/ME+assets).

TLMTA is the ratio of Total Liabilities, and the sum of company Market Equity to Total Liabilities.

TLTA is the ratio between company Total Liabilities and company Total Assets(adjusted).

RSIZE is the logarithm of the ratio of company Market Equity to the S& P500 Market Value.

WC is the company Working Capital over total assets.

Table B.1: Descriptive Statistics

The table reports the summary statistics for all the variables used in the analysis. The sample consists of the intersection of the Compustat, CRSP, and the UCLA- LoPucki Bankruptcy Research Database, over the years 1980-2014. For each variable, column (1) reports the number of observations (firm-year), columns (2)-(3) the mean and standard deviation, columns (4)-(10) the percentile distribution. Panel A refers to the main variables used in our analysis, Panel B to the control variables for the entire sample.

	Obs.	Mean	Std. Dev.	Min	1%	25%	Median	75%	90%	Max
Panel A: Main Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Excess Value	76,389	-0.025	0.647	-2.194	-1.588	-0.412	-0.021	0.325	1.095	3.015
Excess PD	76,389	-0.110	1.257	-3.450	-3.390	-0.894	-0.049	0.715	1.976	2.771
Excess PD (CRI)	16,205	-0.031	0.735	-1.400	-1.370	-0.621	0.000	0.568	1.171	1.399
PD (Estimated - Campbell et al. (2008))	76,389	0.019	0.032	0.000	0.000	0.004	0.008	0.020	0.068	0.770
PD (CRI)	31,089	0.008	0.032	0.000	0.000	0.000	0.001	0.006	0.033	0.870
Default (Y/N)	76,389	0.019	0.137	0.000	0.000	0.000	0.000	0.000	0.000	1.000
Failure (Y/N)	76,389	0.020	0.140	0.000	0.000	0.000	0.000	0.000	0.000	1.000
Numb. Segments	76,389	1.654	1.186	1.000	1.000	1.000	1.000	2.000	4.000	21.000
Panel B: Control Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Size	76,389	5.535	1.811	2.254	2.491	4.136	5.296	6.706	8.925	11.836
Age	76,389	15.534	12.282	0.000	1.000	6.000	12.000	23.000	40.000	65.000
EBITDA	76,389	0.119	0.115	-0.751	-0.300	0.074	0.128	0.181	0.280	0.432
CAPEX	76,389	0.075	0.088	0.000	0.000	0.020	0.048	0.096	0.252	0.658
Sales growth (SG)	76,389	0.139	0.265	-0.418	-0.408	0.000	0.095	0.220	0.667	1.225
Dividends (Y/N)	76,389	0.010	0.021	0.000	0.000	0.000	0.000	0.014	0.044	0.264
Leverage	76,389	0.252	0.183	0.000	0.001	0.101	0.231	0.371	0.597	0.866
LTAT	76,389	0.515	0.197	0.062	0.115	0.371	0.517	0.653	0.845	1.210
CACL	76,389	2.315	1.610	0.000	0.000	1.326	1.937	2.816	5.317	14.244
NITA	76,389	0.015	0.128	-2.078	-0.468	-0.001	0.040	0.074	0.132	0.319
TLTA	76,389	0.493	0.197	0.037	0.095	0.346	0.496	0.636	0.824	0.967
EXRET	76,389	-0.008	0.132	-0.586	-0.430	-0.068	-0.005	0.058	0.205	0.418
NIMTA	76,389	0.006	0.120	-1.978	-0.433	0.000	0.030	0.050	0.098	0.360
TLMTA	76,389	0.428	0.247	0.006	0.032	0.226	0.402	0.607	0.878	0.997
EXRETAVG	76,389	-0.014	0.063	-0.353	-0.215	-0.038	-0.012	0.017	0.082	0.220
SIGMA	76,389	0.050	0.060	0.000	0.000	0.012	0.032	0.062	0.178	0.303
CASHMTA	76,389	0.089	0.120	0.000	0.000	0.016	0.047	0.116	0.314	1.204
MB	76,389	2.409	2.974	0.005	0.032	0.908	1.581	2.749	7.153	45.027
PRICE	76,389	19.568	18.457	0.450	0.580	6.625	15.190	25.250	55.000	123.030

Table B.2: Pairwise Correlation

The table reports the pairwise correlation for the main variables of our analysis. The sample consists of the intersection of the Compustat, CRSP, and the UCLA- LoPucki Bankruptcy Research Database, over the years 1980-2014. The symbols * denote statistical significance at the 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
PD	-0.3504*													
CRIPD	-0.2185*	0.2693*												
Default	0.0128*	0.1387*	0.0264*											
Failure	-0.0577*	0.0413*	0.0386*	0.008										
Conglomerate	-0.0595*	0.0446*	0.0444*	0.008	0.9672*									
CFCORR	-0.0375*	-0.1469*	-0.0619*	-0.6220*	-0.005	-0.004								
Numseg.	0.020	-0.1849*	-0.012	-0.1056*	0.001	- 0.003	-0.6186*							
Age	-0.0659*	-0.2016*	0.00	-0.2670*	-0.007	-0.004	0.2711*	0.1897*						
Size	0.2490*	-0.3536*	-0.1199*	-0.1660*	-0.0151*	-0.0123*	0.2060*	0.3410*	0.3373*					
Leverage	-0.0372*	0.1034*	0.1494*	-0.0772*	0.0696*	0.0704*	0.0735*	0.0370*	0.0454*	0.1517*				
EBITDA	0.2294*	-0.3285*	-0.1192*	-0.0247*	-0.0581*	-0.0611*	0.006	0.0468*	0.0342*	0.1418*	-0.0779*			
CAPEX	0.1382*	-0.0241*	-0.0311*	0.0542*	-0.0262*	-0.0264*	-0.0788*	-0.019	-0.1731*	0.0485*	0.0721*	0.2476*		
Sales growth	0.1626*	-0.0275*	-0.0417*	0.0785*	-0.0287*	-0.0313*	-0.0811*	-0.0378*	-0.2726*	0.0123*	-0.0255*	0.1958*	0.2889*	
Dividends	0.1252*	-0.1882*	-0.0300*	-0.1124*	-0.0199*	-0.0189*	0.1075*	0.0951*	0.2172*	0.1306*	-0.1036*	0.2406*	-0.0208*	-0.0870*

* p<0.1

C Appendix: Propensity to diversify

C.1 Survival Probability Estimation

In this section, we provide some extra robustness test of the paper, together with the details of the survival analysis that computed the firm default probability. We start to build our measure of expected default probability (PD) for each firm-year by following Campbell et al. (2008), who elaborate on previous work on survival probability by Shumway (2001) and Chava and Jarrow (2004). First, based on the hazard model:

$$P_{t-1}(Y_{i,t} = 1) = [1 + exp(-a - bx_{i,t-1})]^{-1} \tag{C.1}$$

where Y_{it} is an indicator variable equal to one when the company goes bankrupt at time t. The vector x includes a set of predictive variables affecting firms likelihood to default. As in Campbell et al. (2008), we experiment with two different dependent variables, a narrower one (default) and a broader one (failure), the latter including the default events on bonds, as alternative indicators of financial distress. We estimate the survival model following Equation (C.1) on the Compustat sample. We also estimate a modified version of the model in columns (5) and (6), where we add the conglomerate dummy. It is important to stress that this is not a future probability. In computing the survival probability for each year, it takes into account the stopping time until then, so it does not considers all the default events in the following years. The following table C.1 reports the estimation of the expected default probability as in Campbell et al. (2008).

The table C.1 shows that conglomerates have higher survival probabilities by 8% that focused companies. This suggests that, on average, the coinsurance function dominates over contagion in conglomerates, which is in line with past findings (Hann et al. (2013)).[15] The coefficients of the control variables confirm the findings of Campbell et al. (2008). Larger size, higher income, and higher stock returns are associated with lower default probabilities, while higher leverage and stock volatility are associated with higher default risk.

[15] Santioni et al. (2020) finds that , the probability that a firm belonging to a group survives from 2006 to 2013 is about 56%, compared to about 50% for unaffiliated firms, in Italy.

Table C.1: Survival Analysis

The table reports the estimates of the default probabilities according to the model of Campbell et al. (2008), where the dependent variable is an indicator variable equal to one when the company goes bankrupt, or fail, in t, and X a vector of variables listed in the table. Columns (1)-(4) report different versions of the survival model, while in columns (5) and (6) we add the dummy conglomerate to the baseline estimation. The estimates are computed with robust standard errors. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. Var.:	Default	Failure	Default	Failure	Default	Failure
	(1)	(2)	(3)	(4)	(5)	(6)
Conglomerate					-0.302***	-0.298***
					(-4.798)	(-4.859)
NITA	-1.893***	-1.881***				
	(-18.352)	(-18.244)				
NIMTAAVG			-1.639***	-1.621***	-1.638***	-1.619***
			(-13.141)	(-13.176)	(-13.053)	(-13.083)
TLTA	4.289***	4.417***				
	(29.48)	(30.48)				
TLMTA			3.251***	3.410***	3.321***	3.479***
			(27.46)	(28.88)	(27.73)	(29.16)
EXRET	-1.331***	-1.356***				
	(-7.268)	(-7.536)				
EXRETAVG			-3.315***	-3.255***	-3.307***	-3.248***
			(-8.966)	(-8.868)	(-8.958)	(-8.860)
SIGMA	2.937***	2.954***	1.437***	1.451***	1.391***	1.405***
	(8.56)	(8.76)	(4.26)	(4.38)	(4.12)	(4.25)
RSIZE	-0.204***	-0.196***	(0.028)	(0.023)	(0.022)	(0.018)
	(-4.399)	(-4.332)	(-1.083)	(-0.953)	(-0.889)	(-0.751)
CASHMTA			-1.968***	-1.857***	-1.959***	-1.849***
			(-6.836)	(-6.773)	(-6.854)	(-6.790)
MB			0.059***	0.059***	0.058***	0.057***
			(7.90)	(7.85)	(7.69)	(7.65)
PRICE			-0.035***	-0.035***	-0.033***	-0.033***
			(-8.000)	(-8.266)	(-7.704)	(-7.953)
Constant	-6.521***	-6.564***	-5.431***	-5.497***	-5.404***	-5.471***
	(-62.541)	(-62.963)	(-51.946)	(-53.091)	(-51.771)	(-52.922)
N	87,427	87,427	87,427	87,427	87,427	87,427

C.2 Propensity to diversify

In this section, we report the details of the estimation of the propensity to diversity by using different models (baseline and enhanced), for the difference-in-difference propensity score matching estimation, where the treated firms are those that switch from focused to conglomerate, and the control firms are focused firms that never change their status. The models are based on Villalonga (2004a), which uses two sets of controls: a set of standard controls that includes firms' assets, EBITDA, CAPEX, industry q and lagged industry-adjusted q, and an enhanced set of controls that also includes firm age, R&D intensity, dummies for major exchange, S&P index inclusion, and firm foreign incorporation.

Variables are defined as follows. EBIT is the ratio of company Earnings to company sales. CAPEX are the firm capital expenditures scaled by firm total sales. The Industry (and Industry-adjusted) q are computed as the median of all focused companies industry tobin q, computed in the same 3 digits SIC Code. The variables S&P, Major Exchange, and Foreign incorporation are indicator variable equal to one when the firm belongs to the S&P index of to a major exchange (NASDAQ, NYSE or AMEX), or the firm has a foreign incorporation. We also control for the firm expenses in research and development (RD, scaled by total assets), and for the fraction of firms that are conglomerate in the same industry (three digits SIC code), and their sales.

The estimation is a propensity score model where the dependent variable is an indicator variable equal to one for firms propensity to diversify, zero otherwise. All models include year effects. The sample of firms switching from focused to conglomerates compose the treated sample, while focused firms compose the control sample. The firms are observed one year before and after the switching, and matched according to the variables reported in Table C.2, Columns (1) and (2). In columns (3) and (4), we report the estimates fromVillalonga (2004a) for comparison purposes.

Specifically, the results confirm that big firms are more prone to diversification, which is the main variable that drives the decision to diversify. Similar to (Villalonga, 2004a), CAPEX has a negative effect on diversification, which confirms the past findings of diversifying firms invetsing less than focused firms in normal times (Glaser and Mueller (2010), (Rajan et al., 2000)). While the industry Tobin-q is negative in the baseline model, it turns insignificant in the augmented model, while is positive and statistically significant in the sample of (Villalonga, 2004a).

This result may reflect the conflicting evidence about the relationship between firms investment and firm market value that has been documented in some works which show that, as investments are positively related to the discrepancy between the market value of installed capital and its replacement cost (Anderson and Garcia-Feijoo, 2006), controlling for the value of firms becomes redundant once the investment factor is added (Fama and French, 2015). Overall, the results confirm the main drivers for the decision to diversify as found from past researchers: the presence of economies of scale, and the firms being mature with less investment opportunity pushing the firms to invest in alternative industries.

Table C.2: Propensity to Diversify

This table reports the propensity score estimation on the subsample of firms that change their status from single to multiple segment firms. The dependent variable is the variable "treated", an indicator variable equal to one if the company change status from single to multi-segment firms, zero for focused firms. Columns (1)-(2) report the probit estimates from two different models for the propensity to diversify of the firms in our sample. The sample in columns (1) and (2) includes all the firms that change their status from one to multiple segment with data one year before, and one year after the change of status, plus focused firms, over the years 1980 - 2014. for comparison purposes, in columns (3)-(4) we report the same models estimates from Villalonga (2004a) on a sample period ranging from 1976 to 1997. The model controls for a vector of company and industry characteristics (listed in the table), including year fixed effects. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Model 1 (1)	Model2 (2)	Model 1(V) (3)	Model 2 (V) (4)
Log of assets	0.299***	0.209***	0.132***	0.223***
	(13.316)	(7.599)	(6.640)	(6.370)
EBIT/sales	0.277	0.352	-1.163***	-1.910***
	(1.076)	(1.042)	(-2.630)	(-3.190)
Capex/sales	-0.161***	-0.101**	-0.145	-0.133
	(-3.435)	(-1.989)	(0.680)	(-0.470)
Industry q (t-1)	-0.093***	-0.060	0.079***	0.108***
	(-2.626)	(-1.456)	(2.820)	(3.450)
Industry-adjusted q (t-1)	-0.063*	-0.030	-0.092	0.045
	(-1.854)	(-0.824)	(-1.650)	(0.810)
S&P		0.034		-0.196
		(0.337)		(-1.400)
Major exchange		0.000		-0.070
		(0.098)		(-0.066)
Dividends paid		0.100		-0.283
		(0.747)		(-1.240)
Foreign incorporation		-0.475		0.026
		(-0.224)		(0.280)
RD/assets		-0.130		2.301*
		(-0.195)		(1.800)
Log of age		0.552***		0.003
		(11.262)		(0.030)
Fraction diversified firms in the industry		0.268		1.098***
		(0.869)		(4.120)
Fraction sales of diversified firms in the industry		-0.077		0.44***
		(-0.346)		(2.120)
Year FE	Yes	Yes	Yes	Yes
Pseudo R-squared	0.098	0.128	0.030	0.100
N	27,695	27,695	24,689	22,527

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.global/content/working-papers