Survival and Value: the Conglomerate Case

DOI: 10.1111/eufm.12442

ORIGINAL ARTICLE

EUROPEAN FINANCIAL MANAGEMENT WILEY

Survival and value: The conglomerate case

Michela Altieri[1] | **Giovanna Nicodano**[2,3,4]

[1]Department of Business and Management, Luiss Guido Carli, Rome, Italy

[2]Esomas Department and Collegio Carlo Alberto, University of Torino, Torino, Italy

[3]CEPR, London, UK

[4]ECGI, Brussels, Belgium

Correspondence

Giovanna Nicodano, Esomas Department and Collegio Carlo Alberto, University of Torino, CEPR and ECGI, Piazza Vincenzo Arbarello, 8, 10122 Torino, Italy.

Email: giovanna.nicodano@unito.it

Funding information

Ministero dell'Università e della Ricerca, Grant/Award Number: Dipartimenti d'Eccellenza

Abstract

This paper investigates the relationship between default probability and value when there is a selection bias due to missing controls for firm heterogeneous likelihood to survive in the sample. Our model delivers the following implications for the conglomerate case: (a) the sample conglomerate value increases in their default probability, (b) the sample conglomerate discount falls together with their excess default probability with respect to focused companies, (c) both effects disappear when the analyst controls for survival probability. The data support the presence of a selection bias distorting downwards the relative value of sample firms with higher survival probability.

KEYWORDS

coinsurance, conglomerate, default probability, firm value, survivorship bias

JEL CLASSIFICATION

G34, G14

We thank the Editor and gratefully acknowledge the useful suggestions by an anonymous referee. We also thank 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 European Financial Management Conference 2022, the 38th International Conference of the French Finance Association (AFFI), the 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 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

By diversifying across industries, conglomerates increase their survival probability when coinsurance across their component units exceeds potential contagion effects. In turn, higher survival is predicted to increase expected value (Banal-Estañol et al., 2013; Lewellen, 1971). When comparing conglomerates and focused companies in sample data, however, higher survival is negatively associated with conglomerate value (Borghesi et al., 2007) even if coinsurance does reduce the survey-based expected cost of capital (see Hann et al., 2013).

This paper proposes a new solution to this puzzle, based on the observation (Brown et al., 1995) that all empirical finance works condition on firm securities surviving in the sample. This gives rise to a selection (or survivorship) bias, that is a distortion in sample values originating precisely from the different survival of firms. According to our model, the prediction that higher survival probability increases firms' expected value should be observed in (samples that are representative of) the population. However, the typical sample of surviving companies is truncated in proportion to firm default probability and is therefore not representative of the firm population. The consequent selection bias results in a positive association between value and default probability in the sample firms.

To grasp the working of the selection bias, let two types of conglomerates, X and Y, have the same expected values at t_0, when they are born. Then, at t_1, the two sets will have the same average price in the sample of survivors only if they have the same default probability. If, instead, X firms have lower default probability than Y firms, then the X survivors will display a lower average price than the Y survivors—as more Y firms will have defaulted. That is, Y survivors are drawn from a sample with higher average value than X survivors. Our model of the survivorship bias formalizes this intuition, deriving a new set of restrictions on the cross-section of conglomerates implied by the selection bias. In line with such implications, we find that the bias increases conglomerate value in proportion to their default probability in US sample data. Our model also considers the conglomerate discount, a long-debated value puzzle that is observed in several data samples, in a new light. To see how, let X be conglomerates and Y be their focused industry peers in the above example. At t_1, the latter will display a premium, being extracted from a sample with a larger left truncation than the conglomerates. In other words, according to our theory, the conglomerate discount due to the survivorship bias will be directly proportional to the excess survival probability of conglomerates relative to focused peers. This pattern should be detectable only in samples that are truncated according to survival probability when controls for differential survival are missing. It should not be present, instead, at the time of conglomerate formation when the conglomerate sample is representative of the population. Following the example, we should then observe the same value for X and Y since there is no selection bias. Furthermore, the covariation between the conglomerate value (discount) and survival probability should be nonnegative, consistent with the implications of previous theories of firm coinsurance (Banal-Estañol et al., 2013; Lewellen, 1971).

Our survivorship bias model delivers a set of implications that we bring to the data on a sample of US Compustat firms. The first implication is that, due to the selection bias, the excess conglomerate value is negative when the excess conglomerate default probability is also negative. In the data, we measure firm excess value relative to focused industry peers as in the literature on the conglomerate discount. We similarly construct the firm excess default probability, based on the default probability estimates computed according to the survival analysis in Campbell et al. (2008). This new measure captures the ability to survive of conglomerates with respect a portfolio of similar focused companies. In line with the first implication, excess value, and excess default probability are respectively equal to −6% and −24% in the raw data, turning to −9% and −6.5% when we control for firm characteristics (other than differential survival) in the regression.

The second implication of the model concerns cross-sectional covariation of value and survival within the conglomerate sample. Due to the selection bias, the conglomerate (excess) value should be lower when the conglomerate survival probability is higher. To test this implication, we perform a quantile regression and examine the distribution of the excess value conditional on survival probability. Consistent with this implication, conglomerates show a severe discount (10%) in the highest survival probability quantile, while there is no discount in the bottom quantile.

According to the third implication, the sample market value of surviving firms coincides with their bias-free value when the analyst appropriately controls for their heterogeneous survival probability. To test this implication, we repeat our regression analysis after adding to the usual set of controls the excess default probability, instrumented to address its endogenous relationship with firm value. In the spirit of Angrist et al. (1996), we use the 6-year-lagged excess default probability to control for differential default probability. We find that the conglomerate discount drops to zero in the overall sample once controlling for firms' ability to survive.

These results stand several refinements, including different measures of conglomeration, the addition of firm fixed effects, and alternative metrics of firm survival probability. Throughout this analysis, we also control for firm age as it has been observed that the discount is smaller for older firms with fewer growth opportunities (Borghesi et al., 2007). Furthermore, a quantile regression shows that bias-free conglomerate excess value varies from 1.5% for above-the-media-survival to zero in the bottom quartile of survival probability. This result is in line with the theoretical predictions, in Banal-Estañol et al. (2013) and Lewellen (1971), that higher survival probability is associated with higher expected value in a bias-free setting.

Finally, we can shed light on the fourth implication of the model regarding the bias-free valuation when conglomerates are just born out of focused firms. The negative relationship between the conglomerate (excess) value and their (excess) survival probability should disappear. Moreover, the discount should be lower at the time of conglomerate formation than later on, when sample selection biases relative values. For instance, if there is no difference in expected values at $t=0$ (as in our example above) and a discount is observed in the overall sample, the discount should disappear in the sample of newly formed conglomerates. We apply the longitudinal approach used in Lang and Stulz (1994) and Graham et al. (2002) to study the discount when firms start to diversify. The conglomerate discount is both not statistically different from zero and insensitive to variation in survival probability across newly formed conglomerates. Summing up, four different implications support the presence of a survivorship bias suggesting that differences in sample prices of firms with heterogeneous survival probability do not reflect the difference in their population values.

The rest of the paper proceeds as follows. Section 1.1 reviews some related literature on the conglomerate discount, to clarify the difference with our theory of selection-biased firm value. Section 2 presents our model of firm value. Section 3 determines the cum-bias market value of firms that survive in the sample. Section 4 investigates the empirical relation between value and default probability. Conclusions follow. Appendix A defines all the variables used in the empirical analysis, along with descriptive statistics (Table A1).[1]

[1]Supporting Information: The Online Appendix reports the proofs of propositions, an extension of the baseline model, and the estimation of firms' default probability.

1.1 | Related literature

This paper contributes to the survivorship bias literature. Since the early work of Banz and Breen (1986), finance scholars have been aware of a so-called ex-post-selection bias associated with firm default. One type of selection bias originated from the partial registration of past defaults in early databases (see Chan et al., 1995; Kothari et al., 1995). A second type of bias stems from firm behaviour. For instance, selection distorts downward the default cost estimates based on realized defaults, when firms with higher default costs reduce their leverage thereby increasing their survival (see Glover, 2016). Brown et al. (1995) uncover a third type of selection bias that is intrinsic to empirical finance studies, which necessarily condition on the security surviving in the sample while the population includes securities of defaulted firms. Brown et al. (1995) show that this bias impacts return predictability, the equity premium, and event studies, warning that it has more general implications. Our paper uncovers one more relevant effect of this last type of survivorship bias. The model shows that it impacts the relative values of firms in proportion to their relative survival probability. The empirical part finds support for the model implications in the cross-section of firms, thus complementing earlier results involving the time-series of returns (Brown et al., 1995).

Several types of survival-related biases affect studies of fund performance (Brown et al., 1992; Carhart, 1997). Carpenter and Lynch (1999) investigate how they affect the persistence of fund performance. They in particular show that there are spurious reversals in performance comparisons if the analyst conditions on multiperiod survival when evaluating funds. The reason is that funds that did poorly in the ranking period (the 'losers') must subsequently do well in the evaluation period to survive, while ranking-period 'winners' have room to do poorly. Another distortion, the look-ahead bias, occurs when a performance comparison study relies on data that was not yet known during the time period being studied. Daniel et al. (2009) measure the bias induced by a performance comparison between an index and portfolio strategies using index components implemented before the index composition was known. Our method compares firm values in each year t based on information available in the preceding year. It is therefore free from both multiperiod conditioning and the look-ahead biases. Carpenter and Lynch (1999) also note that the 'attrition', or disappearance from the sample, of poor fund performers creates an additional bias that is distinct from the one arising from both multiperiod conditioning and look-ahead bias. It is precisely such attrition that our paper brings to the foreground with reference to stocks rather than funds. In this latter respect, Daniel et al. (2009) highlight the large exit rate and associated survival bias for stocks over the period 1926–2006 in the CRSP database. They observe that on average 25% of names disappeared after 3.3 years, 75% of names disappeared after 14 years and 95% of names disappeared after 34 years. Our paper points to the differential survivorship bias due to poor-performing focused firms being less likely to be observed in datasets than poor-performing conglomerates.

Our paper sheds new light on the value of coinsurance in mergers. According to previous theories, mergers, on the one hand, allow for coinsurance across units exposed to less-than-perfectly correlated industry shocks (as in Boot & Schmeits, 2000; Lewellen, 1971), thereby increasing survival probability and ex ante expected value. On the other hand, unprofitable units may drag profitable ones into bankruptcy (the contagion case in both Banal-Estañol et al., 2013 and Leland, 2007), thereby reducing survival probability and ex ante expected value. While several empirical studies find a positive role of coinsurance in conglomerates (Cestone et al., 2016; Hann et al., 2013; Kuppuswamy & Villalonga, 2016; Santioni et al., 2019), the direct relationship between survival probability and value is elusive. Adding to previous models of coinsurance, we analyze how the survivorship bias distorts merger value in survivors' samples after measuring firm value conditional on the firms surviving in the sample (at an 'interim' stage). We show that interim expected values paradoxically decrease in survival probability, given ex ante values. Thus, our model is the first to highlight that the survivorship bias confounds the assessment of the benefits of coinsurance. It also explains the disconnect between the value effects of coinsurance predicted by previous theories and the ones observed in survivors' data. Finally, it supports the implications of previous theories in survivorship-bias-free experiments.

These results connect to a recent paper (Noe, 2020) on the properties of unconditional distributions which result in dominance conditioned on selection. Noe (2020) also provides several counterintuitive economic examples where the worse in the population becomes the better in the sample conditioned on selection. Our third proposition shows the presence of such ranking inversion when the value of conglomerates and focused firm are compared conditioned on survival, namely a conglomerate discount in the sample when there is a conglomerate premium in the poulation.

Over the years, several theories and empirical papers have proposed explanations for the conglomerate discount (Berger & Ofek, 1995; Lang & Stulz, 1994, and many others). While our paper uncovers another effect influencing the sample conglomerate value, our survivorship bias theory is distinct from such previous work on the conglomerate discount for several reasons.

First, the level of the discount is a key metric in most previous work whose aim is the determination of conglomerate efficiency relative to focused peers (e.g., Rajan et al., 2000, and many others). On the contrary, our paper aims at identifying a survivorship bias. Therefore, the level of the ex ante discount, which is highly sensitive to the chosen data set (Villalonga, 2004b), is indifferent to our purposes. We study instead the sign of the covariation between value (of the discount) and (excess) survival probability. Second, the case of conglomerates is instrumental to our survivorship bias investigation. On the one hand, it allows us to juxtapose the known value effects of the coinsurance-contagion trade-off (Banal-Estañol et al., 2013) in a bias-free setting to the opposite ones in a new setting characterized by sample selection. On the other hand, we can rely on consolidated empirical methodologies to compute the discount, to control for firm characteristics reflecting ex ante values and to examine newly formed conglomerates. Third, all previous conglomerate discount theories consider samples as representative of the firm population (see Boguth et al., 2022; Maksimovic & Phillips, 2013, and references therein). Due to this key distinction, the survivorship bias theory generates a different set of predictions. For instance, the closest papers highlighting the role of bankruptcy risk in generating the conglomerate discount, Mansi and Reeb (2002) and Glaser and Müller (2010), stress that coinsurance brings about a default risk reduction and an associated transfer from shareholders to bondholders. This theory implies a negative correlation between the discount (when debt is measured at face value) and survival probability, like the survivorship bias theory.

However, only the latter theory delivers the additional implication that the relationship between value and survival probability changes at the ex ante stage with respect to the interim stage. As another example, consider endogenous conglomerate diversification stories (such as Campa & Kedia 2002; Gomes & Livdan, 2004; Graham et al., 2002). They imply that the discount is absent ex ante but present at the interim stage. However, these theories are not concerned with and therefore do not deliver implications as to the sign of the covariation between the discount and conglomerate survival at the ex ante and at the interim stage, like we do. We are not aware of any work on the conglomerate discount offering all implications of the survivorship bias theory.

2 | THE MODEL

The model, based on the setup of Boot and Schmeits (2000), studies the relationship between firm value and survival at both an ex ante stage, when all firms exist, and at an interim stage, where only survivors operate. We will focus on conglomerates, that are composed of operating units, comparing them to a similar portfolio of focused companies running only one unit. To highlight the survivorship bias, we keep a simple set up with no frictions beside, possibly, bankruptcy costs. We allow for heterogeneous survival but rule out other differences such as a unit's profitability, debt needs, and bankruptcy costs.

In Section 2.2 we analyze the ex ante stage. We reproduce known results from the literature on mergers associated with purely financial synergies. Proposition 1 recalls the nonnegative covariation between value and survival probability, within each firm type. It also extends it to the positive relationship between excess value of conglomerates over focused firms (i.e., the discount, if negative) and their excess default probability.

Original results are presented in Section 3, where we determine the (average market) value conditional on firm survival at the interim stage. The survivorship bias arises because the prices of defaulted firms are not available, and the analyst is not accounting for such sample truncation. Proposition 2 proves that the value of surviving firms of a given type exceeds its own ex ante value in proportion to their default probability. By extension, Proposition 3 proves that the discount on surviving conglomerates exceeds the ex ante discount in proportion to their excess survival.

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 honour 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, we determine both the (average) value 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, $K_i \geq 0$. At $t = 1$ the state is realized and firms are liquidated.

One company type (F) is 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. In C, 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 in C, 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.[2]

2.2 | The credit spread and the ex ante expected value

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.

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 the Supporting Information: Online Appendix A.1 proves the ranking of credit spreads across company types. It shows that:

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

[2]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 the Online Appendix we allow unit B to coinsure and contaminate unit A, thereby eliminating asymmetric payoffs.

Appendix A states the cash flow restrictions that support our state space and the derivations of Equations (1)–(3). Such restrictions require that, in state HL, the cash flow of unit A is large enough to support unit B, while in state LH, B's cash flows are too small to support unit A. They allow to reproduce known results in Proposition 1a below, without affecting the main results concerning the survivorship bias in the next section. According to Equation (3), conglomerates thus pay a lower credit spread than focused companies. This is due in part to coinsurance, which reduces the chances of default, as in Hann et al. (2013). In part, it is due to contagion across units 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 $i = \text{A, B}$, 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 $p_B(1 - p_A)$. We find that:

Proposition 1. *The bias-free value: At* t = 0*:*

a. *Expected value*, V, *is nondecreasing 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.*

Proposition 1 shows results concerning a bias-free setting, when the sample of companies is representative of the population. In such setting, the price of diversified and focused companies coincides, as in the intuitive example in the introduction, when there are no bankruptcy costs (see Part (a)). When bankruptcy costs are positive, expected value increases in survival probability for all firm types because higher survival probability saves on bankruptcy costs. This first result is a replica of previous insight from Banal-Estañol et al. (2013), without tax-distortions, and Leland (2007), with tax distortions. Part (b) extends this result showing that there is a conglomerate discount only if conglomerates display higher default probability relative to focused firms.

While it is usually impossible to observe a bias-free setting as the one maintained in Proposition 1, we expect no differential sample survivorship bias between focused units that just became conglomerates and focused units that did not. In line with Proposition 1, we should then observe no discount for newly formed-conglomerates unless they display excess default probability with respect to their focused peers.

Let us remark that Proposition 1 implies no ex ante discount on conglomerates, when bankruptcy costs are absent. It will become clear that the level of the ex ante discount is indifferent, because the survivorship bias implies a different sign of the covariation between value (of the discount) and (excess) survival probability at the ex ante and at the interim stage, for any ex ante discount.

3 | THE SURVIVORSHIP BIAS

This section contains new insights regarding the consequences of the survivorship bias. We first determine the market value of companies that survive into the sample at the interim stage, when the state is still unknown to market participants. At this stage, the sample of listed companies no longer coincides with the population of companies at $t = 0$. Market values are therefore equal to the expected values conditional on company survival, for each company type. The following proposition summarizes our finding, concerning the relationship between the market values of each company type and their survival probabilities ($p_j^{Sur} = p_j$ for $j = A, B$ and $p_C^{Sur} = p_A$):

Proposition 2. *The biased firm value: 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 default 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. It directly implies the following measure of the (proportional) survivorship bias, for each firm $j = A, B, C$:

$$\frac{MV_j - V_j}{V_j} = \frac{\left(1 - p_j^{(Sur)}\right)}{p_j^{(Sur)}}. \tag{8}$$

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. While this paper deals with the conglomerate case, Equations (7) and (8) hold for any firm type, for example, industry, size, book-to-market, and so on. Before proceeding, let us note that Equation (8) reminds of the expression for the mean observed equity premium conditional on survival in Brown et al. (1995, p. 861) despite the different model set up. In their model, that focuses on the time series of the stock market returns, $p_j^{(Sur)}$ is the probability that the market will survive in the very long run. However, Julliard and Ghosh (2012), and Li and Xu (2002) argue that an equity market implosion does not seem likely enough to account for the observed equity premium. By focusing on the cross-section of firms, we move the attention to firm default probability, which is considerably high (Bessembinder, 2018), rather than that of a single market implosion episode.

We can now bring Proposition 2 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. We can therefore state the following proposition:

Proposition 3. *The biased conglomerate discount: 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:*

$$MV_C - MV_F = \pi_A\left[\left(p_C^{Sur}\right)^{-1} - \left(p_A^{Sur}\right)^{-1}\right] + \pi_B\left[\left(p_C^{Sur}\right)^{-1} - \left(p_B^{Sur}\right)^{-1}\right] < 0. \tag{9}$$

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 a sample conglomerate discount is conglomerate excess survival. According to Proposition 1b, the pattern is 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 in samples that are not representative of the population of firms.[3] Proposition 3b is an example of the worse in the population (focused firms in the proposition) becoming the better conditioned on selection (Noe, 2020). While the ranking inversions in Noe (2020) cast doubts on the quality of competitive selection, the inversion in Proposition 3b confounds the assessment of the benefits of coinsurance.

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 survivorship bias in the data.[4] Proposition 3b indicates how far the survivorship bias may disturb our understanding of the value of survival-enhancing diversification based on survivors' samples. We may be misled into believing that diversification is value-destroying when, to the contrary, it increases welfare by containing bankruptcy costs.

Summing up, we model 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 sample average prices and the expected value of the population of companies, that is brought about by a survivorship bias. This bias derives from a sample truncation which is proportional to default probability. To deliver this insight straightforwardly, we rely 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 section.

3.1 | Survival and firm value: From the model to the data

The previous sections show the relationship between firm value and survival probability, with and without a survivorship bias. We have stressed the conglomerate case because the

[3] Had we assumed a discount (premium) at the ex ante stage, the statement would have changed to 'There is a larger discount (smaller premium) in the sample of survivors than in the population' under the same conditions.

[4] These opposite patterns are also present in the extension of the model, presented in the Supporting Information: Online Appendix, 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).

implications of extant theories, that do not consider sample selection, are opposite to the ones predicted in the case of sample selection. In the next sections, we test for the presence of a survivorship bias. Our focus on conglomerates and focused firms will help us measuring the discount and controlling for firm characteristics reflecting ex ante values with a consolidated methodology. We will rely on four model implications.

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 9). To examine this implication in the data, we borrow the definition of excess value from the conglomerate discount literature. We also identify conglomerates as multisegment companies and control for both firm characteristics and firm fixed effects to account for differences in operating profits and bankruptcy costs.[5]

The second implication concerns the sensitivity in the value of surviving conglomerates to their own default probability (see Equation 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. This implication carries over to the sensitivity of excess conglomerate values to their own default probability. 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. Thanks to our focus on conglomerates, it is possible to identify the bias-free moment of conglomerate formation and study the ex ante discount with another established method that accounts for the decision to diversify.

We will 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 & Nicodano, 2014) conditional on debt levels.

4 | EMPIRICAL ANALYSIS

4.1 | Data and sample

Our sample combines several data sources from the years 1980 to 2014. We retrieve information on multisegment companies (i.e., conglomerates) from Compustat-Historical Segments. Previous studies associate each conglomerate segment with similar single-segment companies in the same industry to compute excess values. We follow a similar approach, applying both the matching and the sample selection as in Berger and Ofek (1995) and Lamont and Polk (2002). 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 aggregate firm segment sales above 1% of total firm sales in Compustat. We also drop segments with missing sales and SIC codes; firms operating in other noneconomic activities, such as membership organizations (SIC

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

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 we only have implications for firms with nonzero debt and a nonzero default probability according to the model, we drop 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 multisegment 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 for stock price falling below an acceptable level or insufficient capital (codes 552 and 560, respectively). The third source is the UCLA-LoPucki Bankruptcy Research Database (BRD), which reports bankruptcy filings (both Chapters 7 and 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 1526 default events from 1980 to 2014, which represent 1.82% of total observations and 14% of the firms in the sample.[7]

The combination of CRSP and Compustat datasets does not have a differential impact on focused versus conglomerate firms since both are covered equally. However, when the UCLA-LoPucki Bankruptcy Research Database is used in conjunction with these data sets, there is a possibility of a differential effect as it only reports bankruptcy filings of major public companies with assets of at least 100 million. This can result in a sample selection bias, as conglomerates are typically larger than focused companies (Villalonga, 2004a). To address this issue, we integrated information on default from different related sources, including Compustat, CRSP, and Mergent (bond default), with the UCLA-LoPucki Bankruptcy Research Database. This resulted in a proportion of defaulted firms (1.6%) that is consistent with past evidence (Campbell et al., 2008). We are therefore confident that our sample is not influenced by the combination of datasets.

Finally, we retrieve firm characteristics from Compustat North America data set. 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 Appendix A.

[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 3 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]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 data set 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 3848 companies.

Table 1 reports the number of active firms, conglomerates, defaults, and failures per 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). A raw indicator of sample selection is the comparison between the number of active firms surviving into the sample as of 2014, which is equal to 1386, and the number of defunct companies over the sample years, which is equal to 1526. The Cumulative Distress columns report the number of cumulative events of failure from the beginning of the sample. They show that focused companies go in distress more than conglomerate firms, also considering their respective sample proportions. As of 2014, the number of cumulative defaults is almost three times larger for focused firms than for conglomerates.

The table also reports, for each year, the variation in the number of firms due to mergers, new entries (as in Ramey & 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.

4.2 | Variables and univariate statistics

Following the conglomerate discount literature (see, among others, Berger & Ofek, 1995; Villalonga, 2004a), the firm's excess value is computed, for each year *t*, as the natural logarithm of the ratio between its market value and its imputed value. The firm market value is the market value of firm assets (total assets minus the book value of equity plus the market value of equity). 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. As per our theory, survival has an impact on the imputed values. The latter rely solely on the value of the focused firms that are in the sample at the end of the year *t*. All the focused companies that defaulted before have exited the stock market and, therefore, are not included in the calculation of the industry's median firm value.

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 the survival analysis as in Campbell et al. (2008).[8]

We also compute the imputed survival probability, for each segment of the conglomerate, as the median survival probabilities across all the focused firms in a specific industry (three digits

[8]The results of the survival analysis are in the Online Appendix. It should be noted that the computation of default probabilities is not affected by the look-ahead bias, as described in Daniel et al. (2009). This is because we utilize firm characteristics that were available in the previous year (usually December) to calculate the probability of the firm defaulting at time *t*. From the estimation in Column (3) in Table B.1 of the Online Appendix, we retrieve the survival odds ratios and compute the probability of default for each company and for each year accordingly.

TABLE 1 Number of companies per year.

This table reports 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. Default is an indicator variable equal to one if the firm defaults in a specific year. CONGLOMERATE is an indicator variable that is equal to one if the company engages in industry diversification. 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 (Chapters 7 and 11). The failure events also include firms with financial default or with a D rating. The cumulative default column captures the number of cumulative events of failure from the beginning of the sample. Mergers are the number of mergers, retrieved from Compustat and Thompson Reuters Refinitiv database. 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 & 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 nonfinancial and nonutility firms in the United States, over the years 1980–2014.

Cumulative failures										
Year	**Active firms**	**Conglomerates**	**Default**	**Failures**	**All**	**Conglomerates**	**Focused**	**Mergers**	**New entries**	**Other exists**
1980	2016	1105	23	23	23	13	10	181	0	1462
1981	2042	1097	19	19	42	22	20	177	108	1489
1982	2085	1055	22	22	64	30	34	199	162	1490
1983	2150	1020	23	23	87	41	46	241	261	1494
1984	2271	992	31	31	118	53	65	248	386	1573
1985	2278	932	24	24	142	59	83	272	365	1538
1986	2319	870	24	24	166	65	101	256	478	1538
1987	2476	834	32	32	198	69	129	290	626	1591
1988	2450	769	43	43	241	77	164	239	651	1626
1989	2362	725	56	57	298	87	211	446	612	1374
1990	2364	717	52	53	351	96	255	390	668	1419
1991	2405	711	44	46	397	106	291	441	662	1406
1992	2561	739	30	33	430	118	312	524	781	1457
1993	2813	742	26	26	456	127	329	651	930	1514
1994	3087	742	48	49	505	133	372	800	1086	1574

TABLE 1 (Continued).

					Cumulative failures					
Year	**Active firms**	**Conglomerates**	**Default**	**Failures**	**All**	**Conglomerates**	**Focused**	**Mergers**	**New entries**	**Other exists**
1995	3347	750	47	48	553	138	415	934	1199	1654
1996	3606	752	65	67	620	146	474	1132	1283	1632
1997	3642	728	107	108	728	160	568	1216	1333	1561
1998	3313	1124	134	141	869	198	671	1064	1170	1383
1999	2557	870	110	112	981	240	741	725	999	1122
2000	2327	629	86	93	1074	263	811	551	969	1071
2001	2029	602	64	70	1144	283	861	453	710	879
2002	1859	545	30	37	1181	295	886	428	605	758
2003	1692	509	18	24	1205	301	904	452	574	603
2004	1674	506	22	25	1230	311	919	468	644	547
2005	1643	502	22	23	1253	319	934	484	669	484
2006	1627	504	33	34	1287	329	958	468	692	415
2007	1629	479	45	48	1335	340	995	391	698	397
2008	1519	465	37	39	1374	350	1024	304	549	309
2009	1419	434	25	26	1400	359	1041	339	502	235
2010	1407	429	27	29	1429	366	1063	373	537	164
2011	1359	447	26	27	1456	372	1084	397	581	116
2012	1334	438	20	22	1478	375	1103	393	554	70
2013	1341	441	24	27	1505	384	1121	405	585	32
2014	1386	451	20	21	1526	388	1138	406	555	2
Total	76,389	24,655	1459	1526	1526	388	1138	16,738	23,184	35,979

SIC code). 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. Therefore, segments with bigger sales count the most for determining the average survival ability. We use sales as previous work shows that segment 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.

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 for the 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 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 coinsurance. 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. The table also shows that conglomerates survive (16%) more in the sample, with their excess default probability being 10%–24% lower than their industry peers.

The statistics in Table 2, panel B, show patterns of the excess conglomerate value across survival quantiles that are broadly in line with the second implication of the survivorship bias model. In Columns (3), (6), and (9), we report the differences in excess value 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). The covariation between excess value and excess firm default probability is 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.

A large empirical literature on conglomerates indicates variables that affect the (bias-free) value. In the next section, we will test whether these patterns hold after controlling for such variables.

4.3 | Empirical results

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 contemporaneous excess default probability since firm value and survival probabilities are simultaneously determined. They are thus affected by similar covariates.[9]

[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.

TABLE 2 Univariates.

This table reports the statistics for the main 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 10th, 25th, and 50th percentiles of companies' survival probability, and the statistical *t*-test of average differences between conglomerates and focused firms for each group. 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 (PD) is computed as the natural logarithm of the ratio between a company's probability of default 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. CFCORR is the cross-segments cash flow correlation as computed from Equation (A1) in Appendix A. LEVERAGE is the ratio between total debt and company total assets. SIZE is the natural logarithm of company total assets. Age is the logarithm of number of years from firm IPO. EBITDA is the ratio of company Earnings before Extraordinary Items to company Total Assets. SALES GROWTH is the yearly growth of the ratio of Sales and company Total Assets. CAPEX is the ratio of company Capital Expenditure to company Total Assets. DIVIDEND ratio is the ratio of Dividends to Total Assets. The details of all variables are in Appendix A. 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.

	Focused		**Conglomerates**			
	Mean	**SD**	**Mean**	**SD**	**Diff.**	***t*-stat**
Panel A	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Excess value	−0.006	0.649	−0.066	0.641	−0.060***	(−12.06)
Excess PD (estimated)	−0.033	1.199	−0.271	1.357	−0.238***	(−24.53)
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	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)

(Continues)

TABLE 2 (Continued).

Panel A	Focused Mean (1)	Focused SD (2)	Conglomerates Mean (3)	Conglomerates SD (4)	Diff. (5)	*t*-stat (6)
Obs.	51,734		24,655		76,389	

Panel B: Survival skills quantiles	10% Mean (1)	10% SD (2)	10% Diff (3)	25% Mean (4)	25% SD (5)	25% Diff. (6)	50% Mean (7)	50% SD (8)	50% Diff. (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	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.	8523			12,683			20,194		

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 (9) of our model. We thus start our empirical analysis using the covariates suggested in the literature, as follows:

$$\begin{aligned} \textit{Excess Value}_{i,t} &= \alpha + \beta\ \textit{Conglomerate}_{it-1} + \beta_1\ \textit{EBITDA}_{it-1} + \beta_2\ \textit{Salesgrowth}_{it-1} \\ &\quad + \beta_3\ \textit{Size}_{it-1} + \beta_4\ \textit{CAPEX}_{it-1} + \beta_5\ \textit{Dividends}_{it-1} + \varepsilon_{it}, \\ \textit{Excess Default Prob}_{i,t} &= \alpha + \beta\ \textit{Conglomerate}_{it-1} + \beta_1\ \textit{EBITDA}_{it-1} + \beta_2\ \textit{Salesgrowth}_{it-1} \\ &\quad + \beta_3\ \textit{Size}_{it-1} + \beta_4\ \textit{CAPEX}_{it-1} + \beta_5\ \textit{Dividends}_{it-1} + \beta_6\ \textit{NITA}_{it-1} \\ &\quad + \beta_7\ \textit{CALC}_{it-1} + \varepsilon_{it}, \end{aligned} \tag{10}$$

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 including firm age among the covariates, to control for the life cycle of firms' growth options. 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

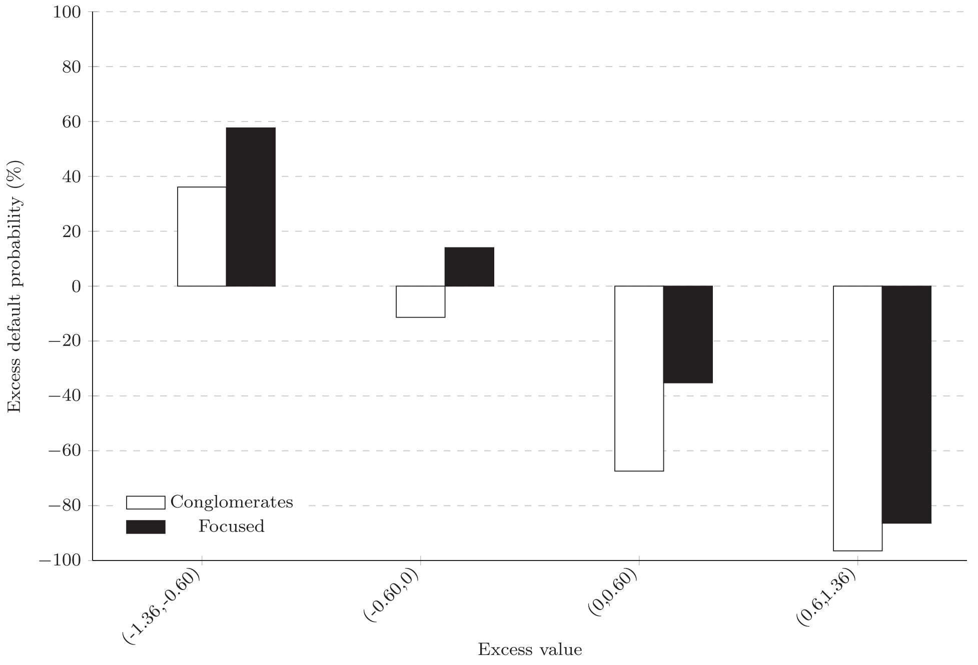

FIGURE 1 Excess default probability by excess values categories. This figure reports the plot of the excess probability of default of conglomerates and focused companies across intervals of the excess value. The excess value is the natural logarithm of the ratio between a company's market value and its 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 Campbell 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.

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 firms 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)).[10] In the reduced sample, the discount increases from 9.4% to 10% with excess survival reaching 14%, indicating that the survivorship bias associated

[10]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 & Livdan, 2004 and Graham et al., 2002) but also contribute to the survivorship bias since low-valuation single-segment companies disappear from the database.

TABLE 3 Excess value and excess default probability: multivariate regression.

This table reports the results of the estimation of the following equation: $y_{i,t} = \alpha + \beta\ Conglomerate_{it} + \Gamma X_{i,t-1} + \varepsilon_{it}$, 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 multisegments. The model controls for a vector of company characteristics (listed in the table), including year and industry (Fama–French 17) fixed effects. The details of variables construction is in Appendix A. 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 (the reduced sample). 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)	(5)	(6)
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)

TABLE 3 (Continued).

	All sample				Reduced sample	
	Excess value		Excess PD		Excess value	Excess PD
	(1)	(2)	(3)	(4)	(5)	(6)
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
Adj-R^2	0.080	0.090	0.371	0.371	0.128	0.412
Obs.	75,393	75,393	75,393	75,393	15,252	15,252

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 & 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 subsamples divided according to 10th, 25th, 50th, and 100th percentiles of company survival probability, as in Equation (10), expecting the conglomerate discount to increase (i.e., the 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 10th and 50th subsamples.

Table 4 reports the results of these regressions. Consistent with the second implication of our model, conglomerates show a severe discount (10% in Column (4)) in the highest survival probability quantile, while there is no discount 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 with lower survival probability display higher value—everything else equal—as they suffer from a smaller survivorship bias. In our robustness section, we perform a similar analysis when using firm-fixed effects to control for firm-and industry-level unobservable variables that can affect the results (Tables 8–10).

4.3.1 | Controlling for the ability to survive in the sample

In our setup, we argue that 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 the ability of firms to survive longer into the sample. We cannot, however, regress current excess value on current excess default probability, because of the simultaneity of both value and default probability, which implies that the regressors and the idiosyncratic error term are correlated and thus coefficients are biased.

To alleviate this concern, we use a lagged variable approach. Specifically, we employ the 6 years lagged excess default probability to control for the effect of survival ability on firms' value. In the spirit of Angrist et al. (1996), we regress firm excess values on the (6 years) lagged excess

TABLE 4 Conglomerate discount and excess default by survival probability.

This table reports the results of the following equation: $pctilereg\ y_{i,t} = \alpha + \beta\ Conglomerate_{it} + \Gamma X_{i,t-1} + \varepsilon_{it}$ 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 10th, 20th, 50th, and 100th, 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 details of variables construction is in Appendix A. 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				**Dep var: Excess Default PD**			
	p (10)	**p (25)**	**p (50)**	**p (100)**	**p (10)**	**p (25)**	**p (50)**	**p (100)**
Survival probability *Q*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(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)

TABLE 4 (Continued).

	Dep var: Excess value				Dep var: Excess Default PD			
	p (10)	**p (25)**	**p (50)**	**p (100)**	**p (10)**	**p (25)**	**p (50)**	**p (100)**
Survival probability *Q*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
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)
χ^2				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
Adj-R^2	0.296	0.252	0.184	0.081	0.181	0.112	0.085	0.202
Obs.	8510	12,785	21,587	43,481	8510	12,785	21,587	43,481

default probability, and use the regression residual as a measure of the excess value that is free from the survivorship bias.

We hypothesize that the excess default probability of 6 years ago, despite being correlated at 50% with current default probability, has a limited correlation with the current value of the firm. We do not add further lags as this condition already requires that both conglomerate and stand-alones remain in the sample for at least 5 years after the IPO. Moreover, Campbell et al. (2008) and Gredil et al. (2022) respectively consider 3 and 5 years as the longest predictive horizon for default probability.

Because some of our explanatory variables are likely to be correlated with unobservable (or unobserved) time-invariant, firm-level omitted variables, we also use firm fixed effects to alleviate this concern. To verify the exclusion restrictions, we employ a sanity check to confirm that 6 years lagged excess default probability has no effect on current firm value.[11] In a second step, we regress firm excess values on the (6 years) lagged excess default probability, and use the residual as a measure of the excess value that is free from the survivorship bias. The results are in Table 5, both on the full and reduced sample. Column 1 shows that, before controlling for any additional variable, the conglomerate discount turns into a small premium (0.1%). The coefficient of the conglomerate dummy is negligible and not statistically different from zero in the next estimations once controlling for all firm characteristics, both in the full and in the reduced sample.

These findings confirm the contraction of the survivorship bias when controlling for differential survival probability. The results support the 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. Once controlling for the covariation between excess value and excess default probability, the conglomerate discount disappears, and possibly turns into a premium.

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 coinsurance. However, they cannot reconcile the coexistence of a lower cost of capital and a lower value of conglomerate firms- suggesting that realized returns contain noise. We take the reasoning one step further, showing that the sample of survivors is not representative of the population.

4.4 | Robustness tests

We perform several additional tests to challenge our baseline findings. First, we bring the lagged approach inside the quantile regression of Table 4, regressing the excess value residuals on four subsamples divided according to 10th, 25th, 50th, and 100th percentiles of company survival probability. Since the dependent variable is now the bias-free excess value, we expect it not to fall in survival probability, in contrast to results in Table 4 and in line with Equations (4)

[11]The results are in the Supporting Information: Online Appendix. We add the 6 years lagged default probability as a regressor in our main Equation (10), expecting its coefficient to be not statistically significant. For comparison, we also add as a regressor the current (simultaneous to the excess value) excess default probability, which we expect being highly significant. We report the results of this model in Table B.2. in the Supporting Information: Online Appendix. In line with our expectations, the coefficient on the 6 years lagged excess default probability is both economically and statistically insignificant. On the contrary (and as expected), the current excess default probability is highly statistically and economically relevant for the current value.

TABLE 5 Controlling for survival ability.

This table reports the results of the estimation of our baseline model after controlling for the excess default probability. We estimate Equation (10) after using as dependent variable the residuals from regressing the excess value on the six periods lagged excess default probability. We estimate the model on the full sample and on a reduced sample (after dropping all firms with corporate events: new firms, mergers, exits for unspecified reasons). 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 variable 'conglomerate' is an indicator variable equal to one if the company is multisegments. The model controls for a vector of company characteristics (listed in the table), including year and firm fixed effects. The details of variables construction is in Appendix A. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Full sample			**Reduced sample**
Dep var: Excess value	**(1)**	**(2)**	**(3)**	**(4)**
Conglomerate	0.008**	0.001	0.003	0.002
	(2.23)	(0.31)	(0.68)	(0.51)
Age			−0.055***	−0.094***
			(−4.469)	(−6.051)
Leverage		0.003	0.000	0.008
		(0.31)	(−0.008)	(0.69)
Assets		0.034***	0.035***	0.034***
		(12.02)	(12.61)	(9.97)
CAPEX		−0.076***	−0.078***	−0.082***
		(−5.094)	(−5.227)	(−4.422)
Sales growth		−0.009**	−0.011***	−0.012***
		(−2.559)	(−2.968)	(−2.594)
Dividends		0.207***	0.213***	0.161*
		(3.01)	(3.08)	(1.84)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adj-R^2	0.511	0.524	0.525	0.520
Obs.	29,358	29,358	29,358	20,949

and (5). Accordingly, results in Table 6 show a conglomerate premium in the upper quantile of survival probability. In other words, there is a conglomerate premium (discount) when coinsurance (contagion) dominates, in line with Proposition 1 and previous theories of firm diversification.

TABLE 6 Quantile regression: controlling for lagged default probability.

This table reports the results of the estimation of our baseline model after correcting for the firms superior ability to survive. We estimate Equation 10 on four subsamples, split according to the 10th, 25th, 50th, and 100th percentiles of companies' survival probability (the latter computed as in Campbell et al. (2008), after correcting our sample for the superior ability to survive by using as dependent variable the residuals from regressing the excess value on the six-periods lagged excess PD. 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 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. The details of variables construction is in Appendix A. In all specifications, the standard errors are clustered at firm level. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	After controlling for six lags excess PD			
	p (10)	p (25)	p (50)	p (100)
Survival probability distribution	(1)	(2)	(3)	(4)
Conglomerate	−0.009	0.002	0.001	0.015***
	(−1.386)	(0.35)	(0.13)	(4.15)
Age	−0.012**	−0.008**	−0.007*	0.004
	(−2.222)	(−2.050)	(−1.856)	(1.22)
Assets	0.012***	0.015***	0.018***	0.027***
	(6.11)	(9.96)	(13.02)	(20.73)
EBITDA	−0.128***	−0.093***	−0.069***	0.066***
	(−5.167)	(−3.868)	(−3.322)	(3.22)
CAPEX	−0.096***	0.012	−0.018	−0.068***
	(−2.748)	(0.51)	(−0.863)	(−3.250)
Sales growth	−0.010	−0.010	−0.008	−0.018***
	(−1.026)	(−1.259)	(−1.211)	(−3.151)
Dividends	−0.104	0.127	0.403***	0.546***
	(−0.597)	(0.911)	(3.898)	(5.789)
Chi2				7.16
Prob χ^2				0.007
Industry FE		Yes	Yes	Yes
Year FE		Yes	Yes	Yes
Adj-R^2	0.050	0.059	0.074	0.184
Obs.	3129	4654	7390	14,185

We then 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 study such ex ante, bias-free discount, 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. We adapt this experiment to determine whether firms that just became conglomerates and display low survival have higher valuations, in contrast to the fourth implication of the model. To this end, 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 operational effects affecting firm value, we should find that the interaction term is significant already at the beginning of the period.

We begin the experiment by identifying 381 firms that transitioned from a focused firm to a conglomerate.[12] 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 1 year before until 1 year after diversification plus the 30,173 single-segment firm-years with data 1 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, as follows:

$$y_{i,t} = \alpha + \beta\ Switchstatus_i + \beta_1\ Switchstatus_i \times after_t + \Gamma X_{i,t-1} + \varepsilon_{it}, \tag{11}$$

where the dependent variable is the excess value of treated (switch status) and control groups, and the treated group is composed of firms that change their status from focused to conglomerate firms (multisegment firms). Focused firms compose the control group. We estimate the difference-in-difference regression as in Villalonga (2004a) over a window of 1 year before/after the change. The vector of controls includes industry (Fama–French 17) and year fixed effects. The results are Table 7, while the details of the variables used in propensity score models are in the Supporting Information: Online Appendix. Columns (1)–(4) of Table 7 report the difference in difference estimation on the treated firms, 1 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 enhanced one. 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.

Consistent with the hypothesis of no survivorship bias at the stage of conglomerate formation, the results confirm that there is no premium associated with low survival conglomerates. In Columns (3) and (4), we see that firms becoming conglomerates have the

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

TABLE 7 Excess value at the ex ante stage of conglomerate formation.

This table reports the results of the estimation of the following equation: $y_{i,t} = \alpha + \beta\ Switchstatus_i + \beta_1\ Switchstatus_i \times after_t + \Gamma X_{i,t-1} + \varepsilon_{it}$, where the dependent variable is the excess value of treated and control groups. The treated group (Switch status) is composed of 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 1 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. The details of variables construction is in Appendix A. 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)
Switch status	−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-R^2	0.094	0.092	0.139	0.141
Obs.	30,554	30,516	30,441	30,441

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 having a low prior survival

TABLE 8 Ex post value correction of switching firms.

This 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 9, 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 of 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 Equation (10) 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 six-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. The details of the variables construction is in Appendix A. 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 for PD			**After correction for PD**		
Dep. var: Excess value	**(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

(Continues)

TABLE 8 (Continued).

	Before correction for PD			After correction for PD		
Dep. var: Excess value	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.104	0.118	0.599	0.153	0.232	0.523
Obs.	34,483	34,483	34,483	25,997	25,997	25,997

probability relative to their focused peers. 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. The coefficient of 'switch status' shows no value gain from shifting status relative to the value of the control group of focused firms that did not switch. Finally, the coefficient of 'low survival' shows a discount of 34% for all low-survival firms (both before/after and switching/not switching), consistent with Proposition 1a.

In the second step, we only keep these firms in our sample (until 2014) to see whether, after the time passing of these firms, patterns become more similar to our baseline results because of the emergence of the survivorship bias.[13] We regress company excess value on a conglomerate dummy, as in Equation (10), before and after the correction for the survivorship bias. Table 8 reports the results. Before the correction, firms that switch from focused to conglomerate 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.

4.4.1 | Other tests

According to several models (see Leland, 2007; Luciano & Nicodano, 2014) coinsurance benefits may prompt the conglomerate to increase leverage well beyond its focused counterparts, increasing default probability, when there is a tax-bankruptcy trade-off. However, conglomerate survival probability is higher than the one of focused companies with the same level of leverage. It follows that, conditional on leverage, the survivorship bias implies the hitherto discussed relationship between value differences and default probability 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 the 10th, 25th, and 50th percentiles of company leverage. We also repeat the same quantile regression analysis conditional on leverage correcting for the survivor probability of firms. Table 9 shows the results. Consistent with the second implication of the survivorship bias model, the

[13]Villalonga (2004b) also shows that the discount appears already after 2 years.

TABLE 9 Excess value by leverage groups.

This table reports the results of the estimations of the following model: *pctilereg* $y_{i,t} = \alpha + \beta\ Conglomerate_{it} + \Gamma X_{i,t-1} + \varepsilon_{it}$ where the dependent variables is the excess value, over the years 1980–2014, regressed within leverage quantiles. In Columns (1)–(3), we perform the estimation in the raw data, while in Columns (4)–(6), we perform the estimation after correcting for the survivorship bias, that is, after regressing the 6 years lagged excess default probability on the excess value and take the residuals as a measure of value free from the survivorship bias. 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 multisegments. The model is performed on four subsamples split according to the 25th, 50th, and 100th 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**		
	p (25)	**p (50)**	**p (100)**	**p (25)**	**p (50)**	**p (100)**
Leverage distribution	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Conglomerate	−0.126***	−0.121***	−0.074***	0.002	0.001	0.001
	(−3.519)	(−6.594)	(−4.736)	(1.58)	(2.81)	(−0.274)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.659	0.673	0.638	0.564	0.554	0.529
Obs.	19,112	19,093	38,184	6142	8161	14,624

conglomerate discount falls from 13% (Column (1)) to 6.5% (Column (4)) when leverage increases, bringing the company closer to distress. After controlling for firm ability to survive (Columns (5)–(8) of Table 7), the conglomerate discount disappears in the high-leverage quantile. It turns into a premium for low leverage quantiles (and therefore low default probability), as expected from our model.

Finally, in Table 10, we also report the estimation of our baseline model with different measures of diversification. In Columns (1)–(3), we report the estimation of the diversification discount before correcting for firms' superior ability of survive by using the number of segments, the coinsurance, and the number of industries as a measure of diversification (all variables are defined in Appendix A). In Columns (4)–(6), we report the same estimation after controlling for the survival ability of firms in the sample. Consistent with our model, the discount disappears. Overall, the robustness tests confirm the baseline implications of our model.

TABLE 10 Excess value and measures of diversification.

This table reports the estimations of Equation (10) 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 explained in Section 4.3.1. 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 nonfinancial, and nonutility 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.

	Before correction for PD			**After correction for PD**		
Dep. var: Excess value	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Number of segments	−0.100***			0.004		
	(−7.450)			(1.03)		
Coinsurance		−0.005			0.003	
		(−0.536)			(1.02)	
Number of industries			−0.028***			0.002
			(−4.687)			(1.05)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.609	0.607	0.608	0.527	0.527	0.527
Obs.	75,393	75,393	75,393	28,927	28,927	28,927

5 | SUMMARY AND CONCLUSION

In this paper, we show how a selection bias spuriously inflates firm value in survivors' samples in proportion to their default probability. While most previous studies of the survivorship bias regard the time series of stock returns, we derive a set of cross-sectional predictions for the changing relationship between value and survival in the population and in truncated samples of firms. We then inspect the conglomerate case, as we can connect with a large body of theories and established empirical methods. However, the set of implications of the survivorship bias theory distinguish it from past conglomerate discount hypotheses.

We exploit the idiosyncratic differences in survival probability, both among conglomerates and between conglomerates and focused industry peers, to study the sign of the correlation between (the differences in) firm survival probability and firm values. Consistent with the survivorship bias hypothesis, this correlation is negative in the sample of survivors: the higher is the (difference in) survival probability, the higher the sample discount, after controlling for firm characteristics. Our empirical analysis shows that such counter-intuitive patterns disappear when we undo the bias. We thus reconcile the bias-free patterns in the data with the prediction of past merger theories that value is nondecreasing in survival probability.

The survivorship bias is likely to confound other cross-sectional comparisons between firms with heterogenous survival. While the implications of our model apply beyond the conglomerate case, we leave it to future empirical work to inspect other firm types. Meanwhile, we should mind the survivorship bias as it may confound the assessment of the value of firm survival.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from Compustat, CRI, CRSP, and UCLA-Lo Pucki. Restrictions apply to the availability of these data, which were used under license for this study. Data are available with the permission of the third parties mentioned above.

ORCID

Michela Altieri http://orcid.org/0000-0003-4735-8475
Giovanna Nicodano http://orcid.org/0000-0003-2098-2331

REFERENCES

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Banal-Estañol, A., Ottaviani, M., & Winton, A. (2013). The flip side of financial synergies: Coinsurance versus risk contamination. *The Review of Financial Studies*, *26*(12), 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*(1), 39–65.

Bessembinder, H. (2018). Do stocks outperform treasury bills? *Journal of Financial Economics*, *129*(3), 440–457.

Boguth, O., Duchin, R., & Mikhail, S. (2022). Dissecting conglomerate valuations. *The Journal of Finance*, *77*(2), 1097–1131.

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., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship bias in performance studies. *The Review of Financial Studies*, *5*(4), 553–580.

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*(4), 1731–1762.

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

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Carpenter, J., & Lynch, A. W. (1999). Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics*, *54*(3), 337–374.

Cestone, G., Fumagalli, C., Kramaz, F., & Pica, G. (2016). Insurance between firms: The role of internal labor markets. *CEPR Discussion Papers*, *27*(11336), 386.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1995). Evaluating the performance of value versus glamour stocks the impact of selection bias. *Journal of Financial Economics*, *38*(3), 269–296.

Daniel, G., Sornette, D., & Wohrmann, P. (2009). Look-ahead benchmark bias in portfolio performance evaluation. *The Journal of Portfolio Management*, *36*(1), 121–130.

Duan, J.-C., Sun, J., & Wang, T. (2012). Multiperiod corporate default prediction: A forward intensity approach. *Journal of Econometrics*, *170*(1), 191–209.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Glaser, M., & Mü ller, S. (2010). Is the diversification discount caused by the book value bias of debt? *Journal of Banking & Finance*, *34*(10), 2307–2317.

Glover, B. (2016). The expected cost of default. *Journal of Financial Economics*, *119*(2), 284–299.

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.

Gredil, O. R., Kapadia, N., & Lee, J. H. (2022). On the information content of credit ratings and market-based measures of default risk. *Journal of Financial Economics*, *146*(1), 172–204.

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*(4), 1705–1745.

Julliard, C., & Ghosh, A. (2012). Can rare events explain the equity premium puzzle? *Review of Financial Studies*, *25*(10), 3037–3076. http://www.jstor.org/stable/41678608

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 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*(4), 905–923.

Lamont, O., & Polk, C. (2002). Does diversification destroy value? Evidence from the industry shocks. *Journal of Financial Economics*, *63*(1), 51–77.

Lang, L., & Stulz, R. (1994). Tobin's Q, corporate diversification, and firm performance. *Journal of Political Economy*, *102*(6), 1248–1280.

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*(2), 765–807.

Lewellen, W. G. (1971). A pure financial rationale for the conglomerate merger. *Journal of Finance*, *26*(2), 521–37.

Li, H., & Xu, Y. (2002). Survival bias and the equity premium puzzle. *The Journal of Finance*, *57*(5), 1981–1995.

Luciano, E., & Nicodano, G. (2014). Guarantees, leverage, and taxes. *Review of Financial Studies*, *27*(9), 2736–2772.

Maksimovic, V., & Phillips, G. (2013). Conglomerate firms, internal capital markets, and the theory of the firm. *Annual Review of Financial Economics*, *1*, 225–244.

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

Noe, T. (2020). Comparing the chosen: Selection bias when selection is competitive. *Journal of Political Economy*, *128*(1), 342–390.

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.

Rosenbaum, P. R. (1984). The consquences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society. Series A (General)*, *147*(5), 656–666. https://doi.org/10.2307/2981697

Santioni, R., Schiantarelli, F., & Strahan, P. E. (2019). Internal capital markets in times of crisis: The benefit of group affiliation. *Review of Finance*, *24*(4), 773–811.

Villalonga, B. (2004a). Diversification discount or premium? New evidence from the business information tracking series. *The Journal of Finance*, *59*(2), 479–506.

Villalonga, B. (2004b). Does diversification cause the diversification discount? *Financial Management*, *33*, 5–27.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114. https://ideas.repec.org/a/eee/ecolet/v104y2009i3p112-114.html

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Altieri, M., & Nicodano, G. (2023). Survival and value: The conglomerate case. *European Financial Management*, 1–39. https://doi.org/10.1111/eufm.12442

APPENDIX A: VARIABLES CONSTRUCTION

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

A.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).

A.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 & French, 1993) factors for each year and industry. Next, we estimate pairwise industry correlations using the previous 5-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:

$$CFCorr_{it(n)} = \sum_{p=1}^{N} \sum_{q=1}^{N} w_{ip(j)} w_{iq(k)} \times Corr_{jk}[t-10, t-1](j,k), \quad \text{(A1)}$$

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 $Corr([t-10, t-1](j,k))$ is the correlation of industry cash flows between industries j and k over the 5-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.

MARKET VALUE is defined as total assets minus the book value of equity plus the market value of equity (computed by multiplying yearly closing price by the number of outstanding shares).

MB (market-to-book) is the ratio between the market value of company equity 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.

A.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 ratio 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 A1 Descriptive statistics.

This 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.

Panel A: Main variables	**Obs. (1)**	**Mean (2)**	**SD (3)**	**Min (4)**	**1% (5)**	**25% (6)**	**Median (7)**	**75% (8)**	**90% (9)**	**Max (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

(Continues)

TABLE A1 (Continued).

Panel B: Control variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
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

See Table A2

TABLE A2 Pairwise correlation.

This 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 symbol * denotes 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*

*$p < 0.1$.