Bankruptcy risk model and empirical tests

Boris Podobnik, Davor Horvatic, Alexander M. Petersen, Branko Urošević, and H. Eugene Stanley

We analyze the size dependence and temporal stability of firm bankruptcy risk in the US economy by applying Zipf scaling techniques. We focus on a single risk factor—the debt-to-asset ratio $R$—in order to study the stability of the Zipf distribution of $R$ over time. We find that the Zipf exponent increases during market crashes, implying that firms go bankrupt with larger values of $R$. Based on the Zipf analysis, we employ Bayes’ theorem and relate the conditional probability that a bankrupt firm has a ratio $R$ with the conditional probability of bankruptcy for a firm with a given $R$ value. For 2,737 bankrupt firms, we demonstrate size dependence in assets change during the bankruptcy proceedings. Prepetition firm assets and petition firm assets follow Zipf distributions but with different exponents, meaning that firms with smaller assets adjust their assets more than firms with larger assets during the bankruptcy process. We compare bankrupt firms with nonbankrupt firms by analyzing the assets and liabilities of two large subsets of the US economy: 2,545 Nasdaq members and 1,680 firms traded on the New York Stock Exchange (NYSE) members. We find that both assets and liabilities follow a Pareto distribution. The finding is not a trivial consequence of the Zipf scaling relationship of firm size quantified by employees—although the market capitalization of Nasdaq stocks follows a Pareto distribution, the same distribution does not describe NYSE stocks. We propose a coupled Simon model that simultaneously evolves both assets and debt with the possibility of bankruptcy, and we also consider the possibility of firm mergers.

We find that book values of assets and debt of the US companies that filed for bankruptcy in the past 20 years follow a Zipf scaling (power-law) distribution. The same is true for the values of assets and debt of nonbankrupt firms comprising the Nasdaq. We focus our attention on a single risk indicator, the debt-to-asset ratio $R$, in order to analyze stability of the scaling exponent or establish cross-over regions. In order to capture Pareto and Zipf laws, the literature has typically focused on a single Simon model (13, 14, 16, 17) describing a single dynamic system which does not interact with others. We model the growth of debt and asset values using two dependent (coupled) Simon models with two parameters only, bankruptcy rate and another parameter controlling debt-to-asset ratio. The Zipf law scaling predictions of the coupled Simon model are consistent with our empirical findings.

Quantitative Methods
Our analysis is closely related to the literature on firm size (19, 20). Analyzing data from the US Census Bureau, ref. 20 reported...
that firm sizes of the US firms follow a Zipf law: The number of firms larger than size $s$ is $s^{-\zeta}$, where $\zeta \approx 1$. The Zipf distribution is found for the distribution of city sizes (21) and the distribution of firm sizes (20, 22).

The cumulative distribution is a simple transformation of the Zipf rank–frequency relation, where the observations $x_i$ are ordered according to rank $r$ from largest ($r \equiv 1$) to smallest. For Pareto-distributed variables $s$ with cumulative distribution $P(s > x) \sim x^{-\zeta}$, the Zipf plot of size $s$ versus rank $r$ exhibits a power-law scaling regime with the scaling exponent $\zeta$, where

$$\zeta = 1/\zeta'.$$  \[2\]

**Results of Analysis**

Fig. 1A shows the Zipf plot for prepetition book value of assets $A_b$. The data are approximately linear in a log-log plot with the exponent

$$\zeta_A = 1.11 \pm 0.01.$$  \[3\]

obtained using the ordinary least-squares regression method. For the US data on firm size (measured by the number of employees), ref. 20 reported the value $\zeta \approx 1$. Hence, prior to filing for bankruptcy protection, the book value of firm assets for companies that later underwent bankruptcy satisfies a scaling relation similar to that in ref. 20. The firms with a rank larger than $\approx 500$ start to deviate from the Zipf law, a result of finite size effects as found in data on firm size (20).

It is known that the market equity of firms that are close to bankruptcy is typically discounted by traders (10, 12). In order to study if those changes are size dependent during the time of bankruptcy, we test whether there is a difference in scaling behavior between prepetition and petition firm assets. Fig. 1B ranks the firm book value of assets $A_b$ and firm debt $D_b$. We find

$$\zeta_b = \zeta_D = 1.44 \pm 0.01.$$  \[4\]

Note that Fig. 1 includes only firms with the largest values of $A_b$ and $A_D$. Thus, the firms with the largest bankruptcy adjustments, with potentially small $A_b$ values, are not necessarily included in Fig. 1. Also, a Zipf law is found for the distribution of total liabilities of bankrupted firms in Japan (23, 24).

We obtain that $\zeta_b > \zeta_A$, a discrepancy that could be of potential practical interest. To clarify this point, if $A_b$ is related to $A_D$ by a constant $A_b/A_D \equiv c$, we would observe $\zeta_b = \zeta_A$. However, we observe an increasing relation $A_b/A_D \propto r^{\zeta_A - \zeta_D}$ with rank $r$, meaning that bankrupt firms with smaller $A_b$ have relative adjustments than do bankrupt firms with larger $A_b$.

Our analysis of bankruptcy probability is, due to data limitation, based on book values. One may argue that a more relevant analysis would be based on market values of assets and liabilities. We now demonstrate that using market instead of book values may in fact lead to similar results. For this purpose, let us consider companies for which we have both market and book value data, namely stocks that comprise the Nasdaq. We begin by finding market capitalization of Nasdaq members for each year from 2002 to 2007. The data are available at Bloomberg L.P. Fig. 2A shows the Zipf plot for market capitalization deflated to 2002 dollar values. We find that the market capitalization versus rank for the largest $\approx 1,000$ companies is well described by a Zipf law with exponent $\zeta_M = 1.1 \pm 0.02$, in agreement with ref. 25.

In Fig. 2B we repeat the Zipf analysis using, this time, book values of both assets and debt for the same Nasdaq stocks. The scaling exponents we observe in Fig. 2B are larger than the exponent observed in Fig. 2A. However, market capitalization is best compared with book value of equity $E \equiv A - D$, rather than assets $A$. In Fig. 2C, we find that $E$ also exhibits Zipf scaling with exponent $\zeta_E = 1.02 \pm 0.01$, which is more similar to $\zeta_M$. Therefore, we find qualitatively similar scaling for the existing Nasdaq companies and for companies before they entered into bankruptcy proceedings.

The probability of bankruptcy $P(R)$ is a natural proxy for firm distress (10). Previous studies analyzed defaults of firms traded at NYSE, American Stock Exchange (AMEX), and Nasdaq (10). In contrast, the majority of firms in our dataset are privately held companies. For bankrupt firms in Fig. 3A we show $P(R|B)$ for values of the debt-to-assets ratio $0 < R < 4$. We truncate data to avoid outliers as in ref. 11. We find $P(R|B)$ is right-skewed with a maximum at $R \approx 1$, and $(R) = 1.4 \pm 1.5$.

Previous studies find that bankruptcy risk of NYSE and AMEX stocks is negatively related to firm size (10). In order to test for firm-size dependence of bankruptcy risk with $R$ as bankruptcy measure, we divide the $R$ values into two subsamples based on their value of $A_b$. In Fig. 3A we demonstrate qualitatively that $R$ is size dependent. The probability density functions (pdfs) for small $A_b$ and large $A_b$ are similar in that they both show peaks at $R \approx 1$. However, firms with smaller assets, as measured by $A_b$, have a larger probability of high debt-to-assets ratios $R$ than firms with large assets $A_b$.

In addition, we test for the size dependence by performing the Mann–Whitney $U$ test, which quantifies the difference between the two populations based on the difference between the asset ranks of the two samples. (The null hypothesis is that the distributions are the same.) Because the test statistics $U$ value $= 5.60$, we reject the null hypothesis thus confirming that $R$ depends on $A_b$ at the $p = 0.05$ confidence level.
conclude that the cumulative distribution of dangerously high
largest 1,000 companies as we find for the assets
financial firms), we find that the Zipf plot exhibits a significant
(see Fig. 1). For
liabilities, follows a Zipf law.

\[ R = D / A \]

In Fig. 3B we analyze the Zipf scaling for large \( R \). We find that
the Zipf plot can be approximated by two power-law regimes. For
\( \approx 300 \) firms with \( 0.8 < R < 3 \) (regime I), we find a power-law
regime with \( \zeta_R = 0.57 \pm 0.02 \). Hence, according to Eq. 2 we
conclude that the cumulative distribution of dangerously high
values of bankrupt firms decreases faster with \( \zeta' \approx 1.72 \) for large
\( R \) than the distribution of firm size (20) and firm assets with \( \zeta' \approx 1 \)
(see Fig. 1). For \( R > 3 \) (7% of all data including predominantly
financial firms), we find that the Zipf plot exhibits a significant
cross-over behavior to a power-law regime with \( \zeta \approx 1.58 \).

The conditional probability \( P(B|R) \) that an existing firm with
debt-to-assets ratio \( R \) will file for bankruptcy protection may be of
significance to rating agencies, creditors, and investors. Accord-
ing to Bayes’s theorem, \( P(B|R) \) depends on \( P(R|B) \) (see Fig. 3),
\( P(B) \), the probability of bankruptcy for existing firms, and \( P(R) \),
the probability of an existing company with leverage ratio \( R \). In
order to estimate \( P(R) \), we use the companies constituting the
Nasdaq in the 3-y period between 2007 and 2009 as a proxy
for existing companies. For this time period, we obtain book value
of each firm’s assets and liabilities (the latter serving as a proxy
for total debt). As a result, we obtain 7,635 \( R \) values with median
value 0.48. For existing Nasdaq members, Fig. 4 shows that the
Zipf plot can be approximated by two power-law regimes, where
regime I with \( 3.5 > R > 0.9 \) yields \( \zeta_R = 0.37 \pm 0.01 \). Note that
regime I is similar to the one we find in Fig. 3B for bankruptcy
data. \( P(B) \) may substantially change during economic crises.
Interestingly, ref. 26 analyzes the debt-to-GDP (gross domestic
product) ratio for countries, in analogy to the debt-to-assets ratio
for existing firms, and calculates a Zipf scaling exponent that is
approximately the same as the scaling exponent calculated here
for existing Nasdaq firms.
We estimate the scaling of \( P(B|R) \) using Bayes’s theorem,
\[
P(B|R) = \frac{P(R|B)P(B)}{P(R)} \approx 0.51P(B)R^{1/z_c - 1/z_R} \approx 0.51P(B)R^{0.95},
\]
where we approximate \( P(R|B) \) and \( P(R) \) with power laws
\(-P(R|B) \sim R^{-(1/z_c + 1)} \Delta R \) and \( P(R) \sim R^{-(1/z_R + 1)} \Delta R \). The value of the relevant exponents calculated for regime I are as follows: \( z_c \approx 0.37 \) (see Fig. 4) and \( z_R \approx 0.57 \) (see Fig. 3B), where \( z_R > z_c \) implies that \( P(B|R) \) increases with firm indebtedness quantified by \( R \). The prefactor 0.51 calculated for the regime I we estimate from the corresponding intercepts in pdfs [see Figs. 3B and 4]. In Fig. 4, we find a pronounced cross-over in the Zipf plot for very large values of the \( R \) ratio.

In order to test whether market crash and global recession have significant effects on the scaling we find in the bankruptcy data, in Fig. 5 we analyze the Zipf scaling of the large \( R \) values for three different 3-y periods. For the period 2004–2006, we find a stable Zipf plot characterized by an exponent \( z_R = 0.50 \pm 0.01 \) close to the value we found in Fig. 3B for all years analyzed. For the period 2001–2003 characterized by the dot-com bubble burst, we find a less pronounced cross-over in the Zipf plot between regime I with exponent \( z_R = 0.58 \pm 0.01 \) and regime II. For the period 2007–2009, we find that the Zipf plot exhibits a significant cross-over behavior between regime I and regime II.

Fig. 5 demonstrates the existence of a relatively stable scaling exponent (between 0.5 and 0.6) in regime I over the 9-y period 2001–2009. However, in times of economic crisis, e.g., the period 2007–2009, the exponent in regime I increases, implying that firms go bankrupt with larger values of \( R \). According to Eq. 5, in times of crisis (\( z_R \approx 0.6 \) \( P(B|R) \propto R/(z_c - 1/z_R) \propto R^{1} \) shifts upward compared to times of relative stability (\( z_R \approx 0.5 \)) when \( P(B|R) \propto R^{0.97} \). A cross-over in scaling exponents may be useful for understanding asset bubbles.

**Model**

Our results complement both the literature on default risk as well as the literature on firm growth. According to a study of US firm dynamics, over 65% of the 500 largest US firms in 1982 no longer existed as independent entities by 1996 (27). To explain how firms develop, expand, and then cease to exist, Jovanovic proposed a theory of selection where the key is firm efficiency; efficient firms grow and survive and the inefficient decline and, eventually, fail (15). Many models have been proposed to model default risk (1, 2, 28–31). One strain of that literature (28) develops structural models of credit risk. In these models, risky debt is modeled within an option-pricing framework where an underlying asset is the value of company assets. Bankruptcy occurs endogenously when the value of company assets is insufficient to cover obligations. In contrast, in reduced form models (2) default is modeled exogenously.

In order to reproduce the Zipf law that holds for bankrupt firms, we propose a coupled Simon model, an extension of the Simon model used in the theory of firm growth (13, 14, 16, 17). Here we couple the evolution of both asset growth and debt growth through debt acquisition which depends on a firms assets, and further impose a bankruptcy condition on a firm’s assets and debt values at any given time.

**Simon Rule for Assets.** The economy begins with one firm at the initial time \( t = 1 \). At each step, a new firm with initial assets \( A \equiv 1 \) is added to the economy. With a probability \( p \), a new firm \( i \) is added to the economy as an individual entity at time \( t_i \). With probability \( 1 - p \), the new firm \( i \) is taken over by an already existing firm. The probability that firm \( i \) is taken over by an existing firm \( j \) is proportional to \( A_j(t)/A_i(t) \). Hence, a larger firm is more likely to acquire a firm than a smaller firm. In this expression, the index \( k \) runs over all of the existing firms at time \( t \). We use the value \( A_j(t) \) to be the proxy for the size of the firm \( j \). Simon found a stationary solution exhibiting power-law scaling, \( P(s > x) \propto x^{-\kappa} \), with exponent \( \kappa = 1/(1 - p) \). For an estimate of \( p \), one can investigate venture data to see how venture capitalists dispose of their companies. Even though data suggest \( p = 0.5 \) (see ref. 32), we use a much smaller value \( p = 0.01 \) in order to reproduce Zipf plot in Eq. 4.

**Simon Rule for Debt.** When a new firm \( i \) is created at time \( t_i \), it is assigned debt \( D_i(t_i) = m \), where \( 0 < m < 1 \). For simplicity, we use a single \( m \) value for all firms. If an existing firm \( j \) acquires the new asset \( A_i \equiv 1 \), then \( A_j(t) = A_i(t - 1) = 1 \), and debt \( D_j(t) = D_i(t - 1) = m \). Hence, a firm with assets \( A_j(t) = N \) has debt \( D_j(t) = mN \), implying that the debt-to-assets ratio \( R = m \) is the same for all firms.

In order to introduce variation in \( R \) ratios across firms, we assume that at each time \( t \), a new debt is created in the economy for some company \( j \), so that \( D_j(t) = D_j(t - 1) = m \). Hence, for each time step, there is a new firm receiving debt \( D_i = m \) in addition to

---

**Fig. 4.** Zipf plot of debt-to-assets ratio \( R \) and rank \( r \) for the existing firms of the Nasdaq members over the last 3 y. For the \( \approx 300 \) ratio values smaller than 3.5 and larger than 0.95 the Zipf plot has exponent 0.37. The same regime we fit with the power-law tail of pdf and obtain 1.54 – 3.6 where the exponent \( \xi' = 3.6 \) agrees (see Eq. 2) with the Zipf exponent \( \xi = 0.37 \).

**Fig. 5.** Zipf plot of debt-to-assets ratio \( R \) versus rank \( r \) for the bankrupt firms for the three different 3-y subperiods. During the last 3 y characterized by recession, the Zipf plot exhibits a cross-over behavior. A smaller cross-over in the Zipf plot also exists for the period 2001–2003 characterized by the dot-com bubble burst.
firm \( j \) receiving one unit of debt, where generally \( i \neq j \). The newly created units of debt are acquired with probability proportional to \( A_i(t) \). Hence, the Simon laws controlling the growth of debt \( D_i(t) \) and the growth of assets \( A_i(t) \) are coupled. In our model, richer firms become more indebted, but also acquire new firms with larger probability.

In Fig. 6A, we perform the numerical simulation of the model by generating 500,000 Monte Carlo time steps. We calculate the Zipf distribution of the debt-to-assets ratio \( R \) for different choices of \( m \). Even though debt and hence \( R \) increases with \( m \), the slope of the Zipf plot for \( R \) versus rank practically does not depend on the value of \( m \). Unless stated otherwise, in other simulations we set \( m = 0.5 \).

Following ref. 33, we consider the continuous-time version of our discrete-time model. In this case, \( D_i(t) \) and \( A_i(t) \) are continuous real-valued functions of time. Further, we assume that the rate at which \( D_i(t) \) changes is in proportion to the assets size \( A_i(t) \). Hence, following this assumption, \( D_i(t) = (1 + m)A_i(t) \) because of the acquisition of additional debt. Therefore, because \( A_i(t) = t/1 \) (33), then \( D_i(t) = (1 + m)t/1 \). The cumulative probability that a firm has debt size \( D_i(t) \) smaller than \( D \) is, therefore, \( \Pr(D_i(t) < D) = P(t) > (1 + m)t/D \). In the Simon model we add new firms at equal time intervals. Thus, each value \( t \) is realized with a constant probability \( P(t) = 1/t \). It follows that

\[
\Pr(t > (1 + m)t/D) = 1 - (1 + m)/D = 1 - (1 + m)/D.
\]

Hence, Eq. 6 should be considered as the Zipf law for debt in the case when there is no possibility of bankruptcy (see Eq. 3).

**Firm Bankruptcy.** Up to now, debt has been modeled as riskless. We now introduce bankruptcy into the coupled Simon model. We assume that for each firm there is a likelihood of bankruptcy, which depends on the volatile firm asset value (28). In order to be consistent with our empirical findings, we assume that the firm \( j \) that was created at time \( t_i \) files for bankruptcy with probability \( qR^{5/3} \) (see Eq. 6), where \( q \) is the bankruptcy rate parameter, related to \( P(B) \) in Eq. 6. In the hazard model, the hazard rate is the probability of bankruptcy as of time \( t \), conditional upon having survived until time \( t \) (11). In our model, once firm \( j \) files for bankruptcy, part of its debt is lost (restructured) and the firm starts anew with debt equal to \( D_j = mA_j \). We do not assume a merger or a liquidation and a firm’s probability of failure does not depend on its age (11). Besides bankruptcy, a firm may leave an industry through merger and voluntary liquidation (9).

Next we perform 500,000 Monte Carlo time steps for the model with the possibility of bankruptcy. Fig. 6B presents Zipf distribution for firm asset and debt values for all of the existing firms. Each of these distributions is in agreement with the Zipf law and Eq. 6. In Fig. 6C, for the subset of bankrupt companies, we show the Zipf distribution for \( R \) using three different values of the bankruptcy rate \( q \). Note that \( q \) is supposed to be small. Namely, with \( q = 10^{-1} \) and with 500,000 time steps representing 1 y, 500,000y represents a probability per year that a company files for bankruptcy during a period of 1 yr, no.05 in our case. Our result for the annual probability of bankruptcy should be compared with the average default rate no.04, calculated in the period 1985–2007 (34). We see that model predictions approximately correspond to the empirical findings.

Our model can be extended in different ways, including mergers between firms. First, although the Simon model assumes that, at each time increment a new unit is added, we can assume that the number of new units grows as a power law \( p^r \) (35). By using a continuous-time version of a discrete-time model, we obtain

\[
\Pr(D_i(t) < D) = \int P(t) > (1 + m)t/D = 1 - \left(\frac{1}{1 + m}\right)^{1/q},
\]

where we use \( P(t) > t_0 = \int \frac{dt}{t} \). Second, Jovanovic and Rousseau (32) found that mergers contribute more to firm growth than when a firm takes over a small new entrant. In order to incorporate mergers into the Simon model, we assume that at each time \( t \), a single merger between a pair of firms occurs with probability \( p' \), where two firms are randomly chosen. Ref. 36 reported that, in more than two-thirds of all mergers since 1973, the Tobin Q value of the acquisition firm exceeded the Tobin Q value of the target firm, where \( Q \) is Tobin’s ratio similarly defined as \( D \) ratio in Eq. 1. To this end, we assume that if \( A_j > A_i \) when a merger occurs, \( A_j = A_j + A_i \) and \( A_i = 0 \). Thus, the more-rich firm \( j \) buys the less-rich firm \( i \) receiving one unit of debt, where generally \( i \neq j \).
results reveal a discrepancy in scaling of market capitalization of NYSE stocks follows a Pareto law with exponents that are slightly scaling for market capitalization (see Fig. 8). Companies traded at NYSE, we do not find similar power-law properties are not trivial consequences of the scaling because, for companies with book value of assets, liabilities, and equity for the stocks traded ... of firms will merge. With increasing $p'$, the Zipf exponent $\zeta$ slowly decreases. Note that, with 1 million time steps, if $p' = 0.5p$, and with $p = 0.01$, then approximately 5,000 mergers occur.

Before concluding, we note that market capitalization as well as book value of assets, liabilities, and equity for the stocks traded at Nasdaq exhibit Pareto scaling properties. Pareto scaling properties are not trivial consequences of the scaling (20) because, for companies traded at NYSE, we do not find similar power-law scaling for market capitalization (see Fig. 8A) and book value of equity. However, the book value of assets and liabilities for NYSE stocks follows a Pareto law with exponents that are slightly larger than those we find for Nasdaq stocks (see Fig. 8B). Our results reveal a discrepancy in scaling of market capitalization (e.g., Nasdaq and NYSE).

ACKNOWLEDGMENTS. We thank Xavier Gabaix for helpful suggestions and the Ministry of Science of Croatia, Serbian Ministry of Science and Technology, and the National Science Foundation for financial support.