Recovery rates: Uncertainty certainly matters

Paolo Gambetti[a,*], Geneviève Gauthier[b], Frédéric Vrins[a]

[a]LFIN, Université catholique de Louvain, Louvain-la-Neuve, B-1348, Belgium
[b]GERAD, HEC Montréal, Montréal, H3T 2A7, Canada

Abstract

Previous studies identify default rates as the main *systematic* determinant of bonds' recovery rates. In this work, we revisit this paradigm by investigating the impact of another factor: *economic uncertainty*. We study the influence of the latter and the one of traditional variables on the shape of recovery rates *distributions*, rather than on recovery rates themselves. Based on a wide sample of American default issues and statistical methods tailored to recovery rates data specificities, we show that economic uncertainty is the most important systematic determinant of recovery rates distributions. By contrast, default rate remains a key determinant of the dispersion of these distributions, but not for their means. Taking this evidence into account is critical for the sound implementation of stochastic recovery rates models used by financial institutions for the computation of regulatory capital.

Keywords: Recovery rate, Loss given default, Corporate bond, Credit risk, Uncertainty
JEL classification: G21, G28, G33

1. Introduction

The 2007 financial crisis convinced regulators to put a lot of effort to reinforce risk management practices of banks and financial institutions, especially with respect to credit losses computation and reporting. The enhancement of the European Basel II Advanced Internal Ratings-Based (A-IRB) approach brought by the Basel III regulation and the introduction of the corresponding Advanced Approach in the U.S. are two sound outcomes of this process.

These frameworks set the conditions to allow financial institutions to use internal models for computing capital requirements for credit losses. The computation is based on portfolio

*Corresponding author at: LFIN, Université catholique de Louvain, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. Tel.: +32 10 479 422.
E-mail addresses: paolo.gambetti@uclouvain.be (P. Gambetti), genevieve.gauthier@hec.ca (G. Gauthier), frederic.vrins@uclouvain.be (F. Vrins)

Preprint submitted to Elsevier

risk measures which are dependent on three core factors: exposure at default (EAD), probability of default (PD) and loss given default (LGD). The latter is often represented as (1-RR) where RR stands for recovery rate. Whereas PD and EAD received a lot of attention in the literature, RR is by far the most disregarded factor. This is rather surprising because appearing as a scaling factor in credit losses, it has a huge impact on risk figures. Yet, it is still common practice to set recovery rates at a constant level, typically 35%-40%.

In fact, the actual value of the recovery rate is unknown till the end of the liquidation process, justifying that it should rather be modeled as a random variable. Many empirical evidences support the idea that recovery rates' distributions are driven by *idiosyncratic* factors (i.e. depending on the defaulted securities' characteristics, like e.g. seniority or sector) (Boudreault et al., 2013; Altman and Kalotay, 2014).

However, there is not much literature discussing the *systematic* determinants of bonds' recovery rates. This latter generally recalls the seminal work of Altman et al. (2005) in pointing out that historical default rate impacts the level of recovery rates. In fact, it is the single systematic determinant of bonds' recovery rates identified with precision. More recent studies including macroeconomic variables focus on recovery rates prediction techniques rather than in the identification of precise determinants (Qi and Zhao, 2011; Yao et al., 2015; Nazemi and Fabozzi, 2018).

Our main contribution is twofold. First, we improve the standard methodology to identify the determinants of recovery rates: while previous studies rely mainly on linear regressions to get conditional *means*, we rather explain the recovery rates' conditional *distributions* using specific models tailored to the data, namely the beta regression and its generalization (Ferrari and Cribari-Neto, 2004; Simas et al., 2010). Second, we include *economic uncertainty* (quantified via some specific measures taken from the scientific literature) as another potential systematic determinant, capturing the business cycle effect. An empirical study performed on 1831 American corporate defaults between 1990 and the end of 2013, taken from the Moody's Analytics Default and Recovery Database, leads us to specify earlier findings on the topic. When included in the model, uncertainty actually becomes the most important systematic determinant of the *mean* of recovery rates. Default rates still play an important role, but essentially for explaining their *dispersion*.

Whereas, to the best of our knowledge, this is the first study that sets a link between the literature on recovery rates determinants and the one on economic uncertainty proxies, this connection looks actual quite natural. First, prices of defaulted bonds represent "expected recovery rates" in the sense that they reflect present values of future cash flows, and that the only cash flow of a defaulted bond is the ultimate recovery resulting from the liquidation process[1]. As extensively reported in Acharya et al. (2007) and Altman and Kalotay (2014), this value is strongly influenced by the economic conjuncture, hence by economic uncertainty.

Second, variations in uncertainty are not an endogenous response to the state of the business cycle but rather they are a cause (Gieseck and Largent, 2016). Uncertainty proxies are particularly able to anticipate periods of economic downturns (Ludvigson et al., 2017). The same cannot be claimed about default rates, which are reasonably a consequence of economic downturns. This may further explain why models based on uncertainty proxies perform better than the traditional ones based on default rates.

Eventually, it looks reasonable to assume that in periods of high economic uncertainty, original investors want to quickly sell defaulted securities to specialized investors (like vulture funds), thereby creating sell pressures and prices' drops.

On the top of explaining the systematic factors impacting bonds' recovery rates, our research could be used to set guidelines to be met by banks' internal models under the Basel III Advanced Approach. It also raises regulators' attention about new risk dimensions to be considered for a consistent assessment of credit exposures of financial institutions in

[1]Ultimate recovery rates result from the sale of the defaulted company's assets; the sale of all assets in case of a corporate default implying liquidation (e.g. Chapter 7 default procedure) and only a part of them in case of milder default events which imply a restructuring (e.g. Chapter 11 procedures).

view of mitigating systemic risk.

The paper is structured as follows. Section 2 includes the literature review. Section 3 contains the data description and the preliminary data analysis. Section 4 presents the methodology. Section 5 includes the main results and their discussion. Section 6 contains the conclusion and proposals for further research.

2. Literature review

This paper builds on two different streams of literature which we eventually connect: the one about the determinants of bonds' recovery rates and the one about proxies of economic uncertainty. Let us now review both of them.

2.1. Determinants of recovery rates

The first comprehensive studies on the determinants of bonds' recovery rates appeared in the late 90s and their number increased together with data availability. These empirical studies, developed by both academia and rating agencies, pointed out two main findings: (i) recovery rates exhibit large cross-sectional variations and (ii) they are time-varying.

(i) Cross-sectional variations.

To be more specific, Altman and Kishore (1996) show for instance that recovery rates are strongly influenced by the seniority of the defaulted bond and the bond issuer's industrial sector. They find that bonds with higher seniority and higher degree of collateralization feature higher recovery rates on average. These effects reasonably represent an anticipation of the so-called *absolute priority rule* which regulates bankruptcy proceedings in the U.S. (11 U.S.C. par. 1129(b)) and states that senior creditors have to be reimbursed before junior creditors in case of a liquidation.

As for industry effects, Altman and Kishore (1996) also find that bonds issued by public utilities and chemistry-related firms present significantly higher recovery rates with respect to bonds issued by firms operating in other sectors, even after they are adjusted for seniority.

Schuermann (2004) reinforces the evidence about higher recovery rates in the utility sector in the period 1970-2003 and also presents similar conclusions for recovery rates on bonds issued by technological and telecommunication firms. The main economic justifications for industry effects relate to asset redeployability considerations as presented in Shleifer and Vishny (1992) and also by Acharya et al. (2007) who use bonds' prices at emergence discounted to the default date. On the one hand, recovery rates should be higher for issuers operating in industrial sectors featuring more tangible assets. On the other hand, recovery rates of bonds issued by firms in sectors that registered an high number of distress periods - and default events - should be lower. This is due to the lower capability of similar firms to absorb the large quantity of assets sold by the defaulted ones. The imbalance between demand and supply of assets of the defaulted firms leads to fire sales in periods of industry distress, which in turn result in lower recovery rates.

Not surprisingly, the severity of the default event is an additional significant determinant of recovery rates. The general evidence is that more severe default procedures lead to lower recovery rates. For example, Bris et al. (2006) show that bankruptcy proceedings implying the liquidation of the defaulted firm - i.e. Chapter 7 procedure in the U.S. - lead to lower recovery rates with respect to the ones that only imply a reorganization, as it is the case in Chapter 11 procedures. Previously, Franks and Torous (1994) reported major differences between recovery rates following Chapter 11 reorganizations and the ones in distressed exchanges, these latter being considerably higher than the firsts. A distressed exchange represents the situation in which the contractual terms of a debt obligation are re-discussed but that do not imply any formal bankruptcy of the debt issuer. Distressed exchanges typically result in diminished financial obligations with respect to the original ones. Further evidence on higher recovery rates resulting from this particular type of default is again reported by Altman and Karlin (2009).

The above-mentioned findings about seniority, industry and default type effects are confirmed by other recent research such as Jankowitsch et al. (2014) which is, to our knowledge, the most complete study on the determinants of bonds' recovery rates developed so far. They define market-based recovery rates as the average trading price of the bonds between the default date and 30-days post default. They bring evidence about the importance of a wide number of security-specific characteristics such as bonds' coupon and maturity and CDS availability. They also show the importance of several firm-specific variables which are justified by structural models of credit risk: the presence of default barriers, of investment and financing covenants and the equity value among others. Most importantly, they report a link between bond-specific measures of liquidity and recovery rates. They find that bonds featuring high illiquidity and transaction costs in the post-default market generally obtain lower recovery rates.

(ii) Time-varying features.

With respect to the time-varying behavior of recovery rates, numerous studies including Frye (2000), Hu and Perraudin (2002), Altman et al. (2005) and Boudreault et al. (2013) acknowledge that recovery rates are strongly dependent on economic conditions or an unobserved credit-cycle, as specifically modeled in Bruche and Gonzlez-Aguado (2010). There is no consensus however on the choice of explicit drivers that need to be used in predictive models for recovery rates, except for default rates and default probabilities. Increasing default rates in a given industry are generally associated to lower recovery rates. The considerations of Shleifer and Vishny (1992) and Acharya et al. (2007) relating to imbalances between demand and supply of physical assets sold by defaulted firms are still used to explain these empirical findings. Altman et al. (2005) also show that demand and supply of distressed securities play a critical role in determining average recovery rates.

As for other proxies for the business cycle, there are mixed evidences. Altman et al. (2005) investigate the predictive power of GDP and its returns and the level of the SP500 and its returns. They find GDP to be relatively low correlated with average recovery rates; the significance improves in the case of GDP returns. When included in a multivariate model together with default rates however, these variables prove to be not significant. The corresponding coefficients also feature a counter-intuitive negative sign. Level and returns of the stock market index are reported to be non-significant in multivariate regressions although they feature an expected positive sign. On the contrary, Mora (2012) finds that real GDP growth relieves the significance from default rates when both are included in the models. They further report the simultaneous significance of the stock market return and default rates.

Although the general conclusion of the literature is that default rates are the most important systematic determinant of recovery rates, we show that it may change if we consider proxies of economic uncertainty, that we now introduce.

2.2. Proxies of economic uncertainty

According to Knight (1921), uncertainty describes a context in which economic agents cannot compute objective forecasts about future outcomes. Concepts such as the ones of *probability* and *expectation* do not apply in an uncertainty framework; even the set of possible outcomes to which probabilities are normally attributed is not observable. It is hence distinguished from the notion of risk.

The interest for the quantification of economic uncertainty is originally linked to real option theory and specifically to firms' decisions in the context of uncertainty. One main idea, carried forward by this stream of literature, is that firms pause their investment and hiring processes in periods of high uncertainty. These choices, in turn, have depressive consequences on the aggregate real output (Bernanke, 1983; McDonald and Siegel, 1986; Dixit and Pindyck, 1994; Bertola and Caballero, 1994; Leahy and Whited, 1996; Bloom et al., 2007). The fact that uncertainty is a determinant of the economic outlook, and most importantly of economic downturns, has also been acknowledged by policymakers and central banks (ECB, 2009; Kose and Terrones, 2012; ECB, 2016).

This evidence further increased the interest for its quantification and gave birth to a specific stream of literature on proxies for time-varying economic uncertainty. For ease of exposition, we divide them in three main categories: (i) survey-based, (ii) news-based and (iii) volatility-based measures.

(i) Survey-based proxies.

Survey-based proxies for uncertainty were originally defined as the degree of dispersion around projections of macroeconomic variables made by professional forecasters. An important reference is represented by Zarnowitz and Lambros (1987). In particular, these authors suggest that uncertainty is better captured by the means of the standard deviations of the predictive probability distributions that professional forecasters attach to their point forecasts rather than the standard deviation of the point forecasts itself, although the two measures are commonly correlated. They find that increases in inflation uncertainty are able to predict lower GDP growth and increases in interest rates in the U.S.

More recent survey-based uncertainty measures build, instead, on the discrepancy in answers to forward-looking questions inquired after enterprises and households. In this respect, Bachmann et al. (2013) provide estimators of time-varying economic uncertainty based on forecast dispersions from business surveys. They analyze the impact of time-varying uncertainty on real activity for both Germany and the U.S.; they find that sudden increases in their proxies have small negative effects on the economic activity for the first country but large and persistent negative effects for the latter. Girardi and Reuter (2017) generalize the approach of Bachmann et al. (2013) and compute uncertainty measures using forecast dispersion indicators from surveys addressed to both consumers and firms from different sectors. Their study brings more evidence about the adverse impact of uncertainty shocks on real activity for the euro area.

(ii) News-based indicators.

News-based indicators of economic uncertainty emerged together with data-mining research, and automatic text-analysis. In particular, Baker et al. (2015) provide a specific proxy for policy-related economic uncertainty, commonly referred as the *economic policy* *uncertainty* index. The methodology for its computation rests on the normalized volume of newspaper articles published in a given month, which contain an exclusive set of expressions referring to economic policy uncertainty. Different specifications of the terms that the text-analysis algorithm has to search for allows the author to propose several other proxies, which they call *categorical* uncertainty measures. They investigate the impact of economic policy uncertainty on financial markets and on several macroeconomic indicators such as aggregate investment, employment and real output for the U.S. and other 11 major economies. They conclude that variations in policy uncertainty indicators are able to predict macroeconomic fluctuations and that high policy uncertainty anticipates periods of economic downturns.

Successively, Alexopoulos and Cohen (2015) introduce a refined version of the main news-based indicator of Baker et al. (2015). A better specification (and an enlargement) of the list of expressions used to proxy economic uncertainty allows the authors to create an algorithm more sensible to the articles' context, which minimizes imprecisions related to word omissions and inclusion errors. When analyzing the impact of uncertainty shocks on economic fluctuations and stock market activity, they reinforce the evidence of previous authors: increments in economic-related uncertainty generally anticipate recessions and high volatility episodes.

(iii) Volatility-based proxies.

Volatility-based proxies of economic uncertainty can be divided in two subgroups: the ones referring to the stock market and the ones that refer to forecast error volatility.

Stock market realized volatility and implied volatility indexes such as VXO and VIX are uncertainty measures belonging to the first subgroup. Leahy and Whited (1996), Bloom et al. (2007) and Bloom (2009) were the first to employ these measures to proxy for firm-level uncertainty. The evidence brought by these authors is that uncertainty shocks based on these indexes are generally associated with rapid decreases in industrial production. They build models which explain the empirical evidence through a real-option channel in line with the one described at the beginning of Section 2.2. They also find that the large value of the *wait and see* real-option in periods of high uncertainty makes firms to be less reactive to economic stimuli. In addition, Bekaert et al. (2013) show that the VIX can be decomposed into an uncertainty and a risk-aversion component and analyze their relationship with monetary policy. They find the first component to be the less influenced by monetary policy and to be the one with the stronger effect on the business cycle.

As for measures of time-varying uncertainty based on forecast error volatility instead, Jurado et al. (2015) represent an important reference. They define uncertainty as the volatility of the purely non-forecastable prediction error for macroeconomic or financial variables. They argue that proxies computed from factor models based on macroeconomic variables are more representative of time-varying economic uncertainty with respect to the ones based on financial variables, being the latter excessively noisy. They find their measure of macroeconomic uncertainty to be more stable than stock market volatility and survey-based measures in spotting periods of high economic uncertainty. Moreover, they contradict the *overshooting* effect of Bloom (2009): macroeconomic uncertainty shocks are associated with a persistent decrease in real output and employment rather than a short-term one.

However, the authors do not discuss a possible causal relationship between uncertainty and economic fluctuations. This question is addressed in the work of Ludvigson et al. (2017). The authors find that financial uncertainty shocks can reasonably be a cause for declines in real activity while variations in macroeconomic uncertainty and in the policy uncertainty indicator of Baker et al. (2015) rather represent endogenous responses to economic fluctuations. Recently, Chulià et al. (2017) developed a daily proxy for financial uncertainty based on a similar methodology with respect to the one of Jurado et al. (2015) and Soojin and Rodrigo (2017) combined the latter approach with the use of survey data.

In this work we rely on various measures (belonging to each of the categories described above) for testing the importance of time-varying uncertainty as a determinant of expected recovery rates distributions. The reason is twofold. First, there is no consensus on which is the best proxy for economic uncertainty since each measure features its own specificities and potential drawbacks (ECB, 2016). The inclusion of different measures allows to draw general conclusions about the impact of uncertainty in determining expected recovery rates. Second, while the literature on recovery rates determinants tells us that expected recovery rates are influenced by business cycle fluctuations, the one on uncertainty proxies reports these latter to be determined and/or accompanied by variations in economic uncertainty. Therefore, we can exploit the predictive power of each proxy while remaining silent about a potential causality relationship between uncertainty and economic fluctuations, a question which goes beyond the scope of this work.

3. Data

3.1. Moody's default & recovery database

Our numerical analyses rely on a wide sample of bonds' recovery rates relating to American default issues and extracted from the Moody's Analytics Default and Recovery Database (Moody's DRD). The sample is composed by 1831 observations spanning a period of 24 years, from January 1990 to December 2013. Recovery rates are expressed as bond prices measured 30 days after the default date, which is declared by the rating agency, and divided by the face value.

We focus on default issues for which we have the following information in addition to the recovery rate: seniority of the defaulted bond, industrial sector of the bond issuer, default type, defaulted amount, coupon level, maturity, presence of backing guarantees different from the bond issuer's assets and default date. From Moody's database we also retrieve American default rates computed at a monthly frequency. In order to obtain a measure as free as possible from rating withdrawal effects, we compute the American default rate as the number of default issues registered in a given month divided by the number of firms followed by Moody's in the same period.

3.2. *Uncertainty measures and other systematic variables*

Measures of time-varying uncertainty are selected among the different types described in section 2.2. With respect to survey-based proxies of uncertainty, we employ an inflation uncertainty measure for United States and a proxy of uncertainty relative to both federal and state/local purchases. These indicators build on the forecasts dispersion computed from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. Data are retrieved from the database managed by the authors of Baker et al. (2015) who used them to create a blended version of their news-based indicator.

With respect to new-based measures of uncertainty, we employ the original economic policy uncertainty index of Baker et al. (2015). The series, available at monthly frequency, are retrieved from the same source of the latter data, which also includes details regarding the methodology for its computation.

The stock market volatility-based measure of uncertainty included in our dataset is represented by the stock market implied volatility index VIX, retrieved from the Chicago Board Options Exchange (CBOE) database at a daily frequency. As for proxies based on forecast error volatility instead, we select the one month horizon *financial uncertainty* indicator of Jurado et al. (2015) and Ludvigson et al. (2017). Data were retrieved from the website of the authors.

Figure 1 depicts the dynamics of these uncertainty proxies which are observable at every default date in our sample. A complete matrix of correlation is instead included in Table 1. The maximum correlation is found between the VIX index and the financial uncertainty indicator, with a value of 0.83. This early finding warns us to use them at the same time for our future analyses. With respect to correlations between other proxies however, we do not report potentially harmful values, the second larger correlation being 0.43, which is found between financial uncertainty and inflation uncertainty.

As for other proxies for the business cycle, we also consider quarterly delinquency rates for commercial and industrial loans extracted from the database managed by the Federal

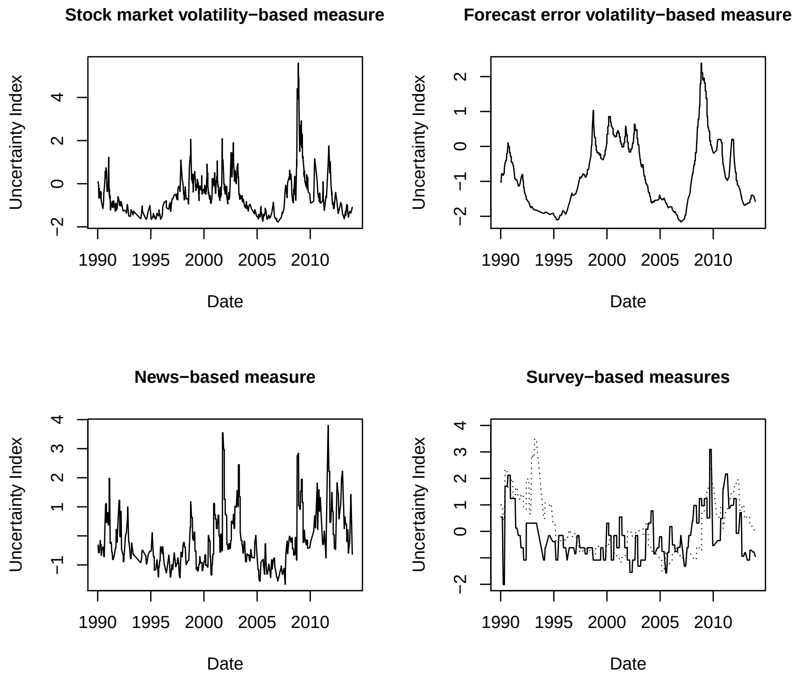

Figure 1: Dynamics of uncertainty measures across default dates. VIX is plotted in the upper-left corner. The financial uncertainty indicator is plotted in the upper-right corner. Economic policy uncertainty indicator is plotted in the lower-left corner. Finally, inflation uncertainty and uncertainty in federal/state/local expenditures are plotted in the lower-right corner as a continuous and a dotted lines respectively.

Reserve Bank of St. Louis, the Federal Reserve Economic Data (FRED). From the same database we also extract the industrial production indicator for United States, quarterly data on the gross domestic product for United States and the NBER based monthly recession indicator. From the database managed by the Chicago Board Options Exchange (CBOE) instead, we retrieve daily data about the level of the stock market index Standard & Poor's 500.

We now proceed with a preliminary data analysis of recovery rates included in our sample.

3.3. Preliminary data analysis of recovery rates data

Table 2 summarizes recovery rates' statistics for our sample set. The average recovery rate is equal to 30.28% and the standard deviation of recovery rates is equal to 25.60%. The lowest recovery rate is essentially equal to zero while the largest observation is exactly one. From the histogram in Figure 2, it can be noticed that the unconditional empirical distribution is highly right-skewed.

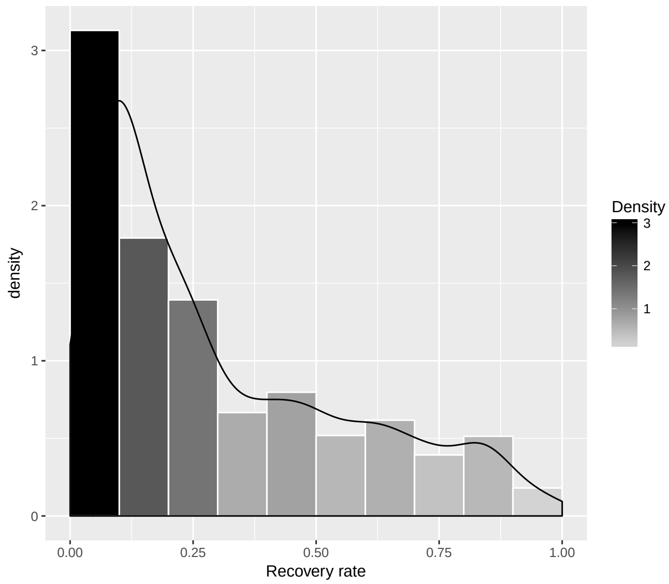

Figure 2: Histogram of recovery rates. Recovery rates relate to American corporate bonds defaulted between January 1990 and December 2013. Darker colors are associated with a higher number of observation in a given bin. The black line represents the empirical density estimated with a Gaussian kernel.

By contrast, one observes very different shapes when considering the empirical distributions of recovery rates conditional on the seniority of the defaulted bonds, the industrial sector of the bonds' issuers and the default type. Summary statistics of these empirical distributions are included in Tables 3, 4 and 5 respectively. Higher levels of seniority and higher degrees of collateralization lead to higher recovery rates, in general. The largest average recovery rate is obtained for defaults on senior secured bonds while the lowest average recovery rate is obtained by junior subordinated bonds. Recovery rates on junior subordinated bonds are also the ones that are less dispersed with a standard deviation value clearly below that of other seniorities. Secured bonds exhibit the largest variation. All recovery rates distributions conditional on seniority are right-skewed except that of senior secured bonds.

As for the empirical distributions of recovery rates conditioned on the industrial sector of the bonds' issuer, bonds issued by banks and technological companies recovered less than the others on average. We also report poor performances for bonds issued by other financial, insurance and real estate companies (FIRE). The largest recovery rates in our sample are obtained for bonds issued by utility companies. The degree of variability seems to be balanced among the different sectors. Moreover, the standard deviation of each conditional distribution is of the same order of magnitude of its mean: even after conditioning on industry, recovery rates still feature a substantial randomness. An exception is made for the distribution of recovery rates on bonds issued by utility companies. This latter is also the only distribution which is left-skewed: all the other distributions are right-skewed with different sizes.

Summary statistics about the empirical distributions of recovery rates conditional upon the default type suggest that more severe default procedures are associated with lower average recovery rates. Recovery rates linked to default types characterized by company reorganizations - such as Chapter 11 and prepackaged Chapter 11 - are evidently lower than the ones associated to default events triggered by missed interest and/or principal payments. The lowest average recovery rates are registered for Chapter 7 liquidation procedures. As for the variability, we observe important differences across conditional distributions. Highest variability is registered for prepackaged Chapter 11 defaults and for defaults triggered by missed principal and interest payments, while we register almost no variability for what concerns liquidation and payment moratorium default events. Finally, the majority of the empirical distributions conditioned on the default type are again right-skewed with different intensities. Only the distributions of recovery rates expected from liquidation procedures and for default events defined as missed principal and interest payments are left-skewed.

A visual comparison among the conditional empirical distributions discussed in this section is included in the online supporting information where we approximated the empirical distributions with Gaussian kernels in order to obtain smoother plots.

4. Methodology

The vast majority of research works published on the determinants of bonds' recovery rates rely on multivariate regression techniques which assume a Gaussian distribution for the dependent variable. This is rather striking because recovery rates are bounded in $[0, 1]$ and present empirical features that are incompatible with this implicit assumption. In fact, as shown in Section 3.3, recovery rates distributions are generally skewed. Recovery rates also feature heteroskedasticity, a typical characteristic of data generated by a stochastic process defined on a closed interval. Gaussian-based regression techniques applied to these data can hence yield deceptive results, such as biased standard errors and p-values. They are also not suitable for prediction purposes.

Standard mappings such as a classical Box-Cox (Box and Cox, 1964) transform may be applied to obtain a normal distribution for the transformed variable or the model's residuals. However, such method does not guarantee that the distribution of the transformed variable is exactly normal. It generally puts some constraints on its first four moments and yields a symmetric distribution (Draper and Cox, 1969). The Manly transformation (Manly, 1976) instead is only able to transform unimodal skewed distributions into normal ones, but cannot deal with U-shaped or multimodal distributions.

Most importantly, by applying one of these transformations as well as other common alternatives such as those described in John and Draper (1980), Bickel and Doksum (1981), Yeo and Johnson (2000) or the simple logit and square-root transformations, the interpretation of the model coefficients is very little intuitive. In fact, regression coefficients have to be interpreted in terms of the mean of the transformed variable instead of directly in terms of the mean of the original one.

In order to address these shortcomings, we base our methodology on two special cases of the general class of beta regression models: the *classical beta regression* (Ferrari and Cribari-Neto, 2004) and the *variable dispersion beta regression* (Smithson and Verkuilen, 2006; Simas et al., 2010) models. These methods have been specifically conceived to deal with the empirical features of percentage data and the model outputs are directly interpretable in terms of the dependent variable.

The regression models used in this work are based on a re-parametrization of the beta probability density function characterized by two shape parameters $\alpha, \beta > 0$:

$$f(y; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha - 1}(1 - y)^{\beta - 1}, \qquad 0 < y < 1, \tag{1}$$

where $\Gamma(\cdot)$ is the Gamma function. Letting

$$\mu = \frac{\alpha}{\alpha + \beta} \in (0, 1) \quad \text{and} \quad \phi = \alpha + \beta > 0 \tag{2}$$

the beta density can be re-written as:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1 - \mu)\phi)} y^{\mu\phi - 1}(1 - y)^{(1-\mu)\phi - 1}, \quad 0 < y < 1\ . \tag{3}$$

We use the notation $Y \sim \mathcal{B}(\mu, \phi)$ to stipulate that the random variable Y has the density given above. This notation allows for an intuitive link between the parameters and the moments. Indeed,

$$E[Y] = \mu \quad \text{and} \quad Var[Y] = \frac{\mu(1 - \mu)}{1 + \phi}. \tag{4}$$

Hence, parameter μ is the mean of the distribution of Y while ϕ can be interpreted as a precision parameter: for fixed μ, the higher its value the lower the variance of Y. In particular, $\phi \to \infty$ corresponds to the Dirac measure $\delta_\mu(x)$ whereas $\phi \to 0$ corresponds to the Bernoulli distribution with mean μ, i.e. the maximum-variance distribution defined on $[0, 1]$.

Let $\{Y_1, Y_2, \ldots, Y_n\}$ be independent random variables, where $Y_i \sim \mathcal{B}(\mu_i, \phi_i)$ for each $i \in \{1, 2, \ldots, n\}$. The variable dispersion beta regression model (as in Simas et al. (2010))

is defined as:

$$g_1(\mu_i) = x_i^T\theta = \eta_{1i}; \qquad g_2(\phi_i) = z_i^T\gamma = \eta_{2i} \tag{5}$$

where $\theta = (\theta_1, ..., \theta_k)^T$ and $\gamma = (\gamma_1, ..., \gamma_h)^T$ are unknown coefficient vectors, $x_i = (x_{i1}, ..., x_{ik})^T$ and $z_i = (z_{i1}, ..., z_{ih})^T$ are regression vectors $(k+h < n)$, η_{1i} and η_{2i} are linear predictors and $g_1(\cdot) : (0, 1) \mapsto \mathbb{R}$ and $g_2(\cdot) : (0, \infty) \mapsto \mathbb{R}$ are strictly increasing and twice-differentiable link functions, hence invertible. The most common link function for μ is the logit $g(\mu) = \log(\frac{\mu}{1-\mu})$. Other possible functions are the probit, log-log, complementary log-log and Cauchy link functions. Common link functions for $g_2(\cdot)$ are instead the logarithmic function $g_2(\phi) = \log(\phi)$, the square-root function and the identity function.

The variable dispersion beta regression models used in this work feature a logit link for $g_1(\cdot)$ and a log link for $g_2(\cdot)$. Parameters μ_i and ϕ_i are hence modeled as:

$$\mu_i = g_1^{-1}(\eta_{1i}) = \frac{e^{x_i^T\theta}}{1 + e^{x_i^T\theta}}; \qquad \phi_i = g_2^{-1}(\eta_{2i}) = e^{z_i^T\gamma} \tag{6}$$

Using equations in (5), we can restate the formula for the variance of the dependent variable as follows:

$$Var[Y_i] = \frac{g_1^{-1}(x_i^T\theta)[1 - g_1^{-1}(x_i^T\theta)]}{1 + g_2^{-1}(z_i^T\gamma)} \tag{7}$$

Considering equation (7), it is evident that the regression model based on this parametrization is naturally heteroskedastic because $Var[Y]$ is a function of μ.

The log-likelihood function for the beta distribution re-parametrized as in equation (3) and based on n observations $\{y_1, y_2, \ldots, y_n\}$, takes the following form:

$$\ell(\theta, \gamma) = \sum_i^n \ell_i(\mu_i, \phi_i) \tag{8}$$

where

$$\ell_i(\mu_i, \phi_i) = \log \Gamma(\phi_i) - \log \Gamma(\mu_i \phi_i) - \log \Gamma((1-\mu_i)\phi_i) + (\mu_i \phi_i - 1) \log y_i \\ + \{(1-\mu_i)\phi_i - 1\} \log(1-y_i). \tag{9}$$

Since $\mu_i = g_1^{-1}(x_i^T \theta)$ and $\phi_i = g_2^{-1}(z_i^T \gamma)$ are functions of the coefficient vectors θ and γ respectively, parameter estimation can be performed by maximum-likelihood. However, the estimation requires the solution of a nonlinear system of the likelihood equations; hence, maximization has to be undertaken with a nonlinear optimization method such as a Newton or quasi-Newton method.

The classical beta regression (Ferrari and Cribari-Neto, 2004) can be considered a simplification of the previous model in which only the parameter $0 < \mu < 1$ is allowed to vary with the observations while $\phi > 0$ is constant. We therefore assume $Y_i \sim \mathcal{B}(\mu_i, \phi)$ for each $i \in \{1, 2, \ldots, n\}$.

Hence, it is enough to substitute equations in (5) with $g_1(\mu_i) = g(\mu_i)$ and $g_2(\phi_i) = \phi_i$ and further impose $\phi_i = \phi$. As before, parameter estimation is performed via numerical optimization of the log-likelihood[2].

An important feature of these methods is that different estimates of μ and ϕ can yield various shapes for the underlying beta distributions: skewed unimodal, J-shapes, inverted J-shapes, U-shapes and even the uniform distribution. Moreover, given the strictly increasing link functions, the sign of the estimated regression coefficients $\hat{\theta}$ for the classical beta regression and $\hat{\theta}$ and $\hat{\gamma}$ for the variable dispersion beta regression can be directly interpreted in terms of the effect of the regressors on the mean and dispersion of the beta distributions generating the observations. However, note that the likelihood functions are only defined in the interval (0, 1) and not at the extremes. Maximum likelihood estimation based on

[2]Details regarding the score functions, likelihood equations and Fisher's information matrices for both models, as well as the form of the other link functions, are included in the online supporting information.

inflated beta distributions (Ospina and Ferrari, 2008) can be implemented in order to take into account of observations at the boundaries[3].

5. Results

As anticipated in Section 4, we do not investigate the impact of the regressors on the observations of recovery rates (as it is done in the majority of the studies about bonds' recovery rates determinants). Rather, we investigate the effect of the regressors on the shape of the beta distributions from which we assume recovery rates are generated.

5.1. Uncertainty as a determinant of expected recovery rates

Table 6 shows the estimated parameters and performance statistics of eight beta regression models with fixed dispersion ϕ as depicted in Section 4. For clarity of exposition, we denote these models as $\mathcal{M}_k^{\mu}$, where $k = \{1, \ldots, 8\}$ indicates the identifier of the model and the superscript μ specifies that we are dealing with fixed dispersion models; only μ varies. The eight predictors all feature a standard logit link for the mean parameter and an identity link for the precision parameter[4].

The purposes of Table 6 are to show how the performance of a model including security-specific characteristics increases as a result of the addition of a systematic determinant, and to compare the performances of alternative models which include only one candidate as systematic determinant of recovery rates. As suggested in Ferrari and Cribari-Neto (2004), we compare the performance of the different models in terms of pseudo-R^2 measures, defined as the squared sample correlation coefficient between the estimated predictor $\hat{\eta}$ and the link-transformed response $g(y)$[5].

[3]Two other possible solutions are adding and subtracting a machine-precision quantity from the extreme observations or to apply a transformation of the form $\tilde{y} = (y - a)/(b - a)$ where a and b denote the sample minimum and maximum (Ferrari and Cribari-Neto, 2004).

[4]The only recovery rate observation exactly equal to 1 is removed from the sample. We hence avoid the implementation of an inflated beta regression model which would require an extended discussion. Moreover, this observation featured an outlier residual even after the subtraction of a machine-precision quantity and it was also marked as a leverage point.

[5]Nested models can also be compared via likelihood ratio tests. Non-nested models can instead be

5.1.1. *Security-specific models augmented for one systematic factor*

Models $\mathcal{M}_1^\mu$ and $\mathcal{M}_2^\mu$ only include security-specific characteristics. Model $\mathcal{M}_1^\mu$ only includes the following categorical variables: seniority of the defaulted bonds, industrial sector of the issuers, default type and presence of additional guarantees different from the bond issuer's asset. Model $\mathcal{M}_2^\mu$ is augmented for taking into account of coupon and maturity. They yield pseudo-R^2 measures of 27% and 28% respectively. The estimates of $\mathcal{M}_2^\mu$ suggest that the mean of recovery rates distributions μ is increasing with the coupon level but decreasing with respect to the maturity of the defaulted bonds, both variables being significant at the 1% level. These results are indeed in line with the findings of Jankowitsch et al. (2014).

Model $\mathcal{M}_3^\mu$ is augmented to include default rates, which for long time have been considered the main systematic determinant of recovery rates. This model yields a pseudo-R^2 around 30% and it confirms past empirical findings in showing an inverse relationship with average recovery rates (as in the seminal work of Altman et al. (2005) and the majority of later research works): higher default rates are able to predict lower means of recovery rates distributions[6].

From $\mathcal{M}_4^\mu$ to $\mathcal{M}_8^\mu$ we substitute default rates with one proxy for time-varying economic uncertainty. Models $\mathcal{M}_4^\mu$ and $\mathcal{M}_5^\mu$ include a volatility-based proxy for uncertainty: the financial uncertainty indicator and the VIX index respectively. Models $\mathcal{M}_6^\mu$ and $\mathcal{M}_7^\mu$ include survey-based proxies: the inflation uncertainty indicator and the indicator relative to federal/state/local expenditures. Model $\mathcal{M}_8^\mu$ includes the news-based economic policy uncertainty index. When analyzing the significance of the coefficients, we first notice that each of these uncertainty proxies is significant in its own model with the exception of the

compared through AIC criterion. We found that both methods lead to the same conclusions of the pseudo-R^2 measure; hence, we base the discussion on this latter, since it remains the simplest global measure of goodness-of-fit.

[6]Here and in what follows, we use the term *predict* because the monthly regressors are included in the models using the values one month before default while daily regressors are measured at the default date. In fact, this structure allows to attribute a casual interpretation to the model results.

one relating to federal, state and local expenditures entering in $\mathcal{M}_7^\mu$. Moreover, all the uncertainty-based models feature a negative sign in the estimated coefficient of the systematic variable: it appears that increasing uncertainty is indeed able to predict lower means for the recovery rates distributions.

Comparing the performance measures of $\mathcal{M}_4^\mu$ to $\mathcal{M}_8^\mu$ with the one of $\mathcal{M}_3^\mu$, we notice that the uncertainty-based models are in general superior to the one based on default rates. For example, $\mathcal{M}_4^\mu$, $\mathcal{M}_5^\mu$ and $\mathcal{M}_8^\mu$ show significant improvements in pseudo-R^2, with measures of 39%, 35% and 32% respectively. Especially for what concerns models built on volatility-based measures of uncertainty - $\mathcal{M}_4^\mu$ and $\mathcal{M}_5^\mu$ - it appears that economic uncertainty is able to explain a higher portion of recovery rates variations than the one explained by default rates. Hence, the results in Table 6 not only show that increasing uncertainty is able to predict lower recovery rates, but they also suggest that uncertainty may have a greater role than default rates as systematic determinant of recovery rates distributions.

5.1.2. Augmented models with other business cycle factors

Table 7 further investigates this question. It includes the results of four beta regressions with fixed dispersion parameter in which we further control for other systematic determinants. Among these control variables we include a monthly polynomial interpolation of quarterly GDP for United States, monthly industrial production returns, a monthly recession indicator, the level of the SP500 and monthly polynomial interpolation of quarterly delinquency rates in commercial and industrial loans.

Given a model $\mathcal{M}_k^\mu$ as above, we denote the corresponding augmented model with other proxies of the business cycle as $\mathcal{M}_{k+}^\mu$. Models which further include all uncertainty proxies are instead denoted by $\mathcal{M}_{k++}^\mu$.

As $\mathcal{M}_{2+}^\mu$ shows, the inclusion of these proxies for the business cycle yields a sensible increase in the model performance. For this model we report a Pseudo-R^2 of 35%, which has to be confronted with the 28% displayed by $\mathcal{M}_2^\mu$ of Table 6 including only security-specific characteristics.

The inclusion of default rates in model $\mathcal{M}^{\mu}_{3+}$ however, does not seem to bring additional information with respect to the one conveyed by the other systematic variables. Although default rates are strongly significant and feature the expected negative sign, the pseudo-R^2 is unaffected. Moreover, given the size of the estimated coefficients, it seems that delinquency rates play a greater role.

Interestingly, $\mathcal{M}^{\mu}_{4++}$ and $\mathcal{M}^{\mu}_{5++}$ - which include proxies of economic uncertainty - come to rather different conclusions. Performance measures increase steadily to 40% and 39% respectively and this gain is not strictly linked to the additional number of variables. In fact, in parallel analyses in which we substitute default rates in $\mathcal{M}^{\mu}_{3+}$ with financial uncertainty, VIX index or news-based economic uncertainty we report Pseudo-R^2 measures of 40%, 38% and 37% respectively[7]. Uncertainty measures indeed convey additional information for explaining recovery rates variations.

Moreover, the evidence about economic uncertainty to be the most important systematic determinant of recovery rates is reinforced. Financial uncertainty is still strongly significant in $\mathcal{M}^{\mu}_{4++}$ when controlling for other proxies of the business cycle. VIX index, the news-based proxy of uncertainty and the survey-based proxy for uncertainty about federal/state/local expenditures are instead significant in $\mathcal{M}^{\mu}_{5++}$. Given the sign of the estimates associated with uncertainty proxies, the main evidence is that increasing uncertainty is able to predict lower means μ of the recovery rates distributions.

Most importantly, significance is relieved from default rates in $\mathcal{M}^{\mu}_{4++}$ and $\mathcal{M}^{\mu}_{5++}$, although the estimated coefficients still feature an intuitive negative sign. Default rates can no longer be considered as key systematic determinants of recovery rates when economic uncertainty is taken into account.

Comparing the performance measures of $\mathcal{M}^{\mu}_{2+}$ and $\mathcal{M}^{\mu}_{3+}$ in Table 7 to the ones of the

[7]Complete results are omitted due to space considerations but they are available in the online supporting information.

lower-dimensional models $\mathcal{M}_4^\mu$ and $\mathcal{M}_5^\mu$ in Table 6, we can also conclude that volatility-based measures of uncertainty are able to explain a greater portion of recovery rates variation with respect to traditional proxies for the business cycle and especially default rates. Hence, economic uncertainty certainly matters and in fact proves to be the most important systematic determinant of expected recovery rates among all the factors considered in this study[8].

Another important question is to determine whether uncertainty is also a systematic determinant of the *dispersion* of the beta distributions from which we assume recovery rates are generated. This question is handled in the next section, where we analyze the determinants of the precision parameter ϕ using a beta regression technique including dispersion covariates.

5.2. Uncertainty and default rates as dispersion determinants

As explained in Simas et al. (2010), the regression vectors x_i and z_i for the mean and precision submodels of variable dispersion beta regressions are not exclusive. This means that the number of parameters to estimate doubles in case a model with the same regressors for μ and ϕ is fitted. To deal with the trade-off between creating an over-parametrized model and understanding what are the most important determinants for the shapes of recovery rates distribution, we perform model selection.

We select the regressors with a backward stepwise model selection algorithm building on the Generalized Akaike Information Criterion (step-GAIC) and based on the last version of Rigby and Stasinopoulos (2008). Variable dispersion beta regressions can in fact be seen as a particular case of the class of generalized additive models for location, scale and shape (GAMLSS) conceived by Rigby and Stasinopoulos (2005). As it is usual for this type of models, we first select variables on the mean submodel; given the best subset of variables

[8]In particular, we find these conclusions to be robust to the specification of different link functions. The complete results are available upon request.

in the mean submodel, we then apply the selection algorithm to the precision submodel. Models in this section are denotes as $\mathcal{M}_k^{\mu,\phi}$ where k is again the id of the model and μ, ϕ specifies that we are dealing with variable dispersion models.

Table 8 includes the results of two generalized beta regressions where the starting models for the step-GAIC procedure are represented by $\mathcal{M}_{4++}^{\mu}$ and $\mathcal{M}_{5++}^{\mu}$ of Table 7. Hence, the starting models for $\mathcal{M}_{4++}^{\mu,\phi}$ and $\mathcal{M}_{5++}^{\mu,\phi}$ include financial uncertainty and the VIX index respectively. As specified in Section 4, the mean submodels feature a logit link while the precision submodels feature a log link.

5.2.1. Robustness of the mean submodel

Again, as for the mean submodels, we notice that uncertainty measures are included both in $\mathcal{M}_{4++}^{\mu,\phi}$ and $\mathcal{M}_{5++}^{\mu,\phi}$. Model $\mathcal{M}_{4++}^{\mu,\phi}$ includes financial uncertainty and the news-based uncertainty proxy, both being very significant and with a negative sign. Model $\mathcal{M}_{5++}^{\mu,\phi}$ selects the VIX and all the other uncertainty measures: all these variables feature the expected negative sign but inflation uncertainty is not statistically significant. Interestingly, we notice that volatility-based proxies of uncertainty are chosen as the *first* systematic variable by the selection algorithm. Financial uncertainty is chosen to be kept in the model already at the third step of the selection algorithm, after the default type and seniority and right before industrial sector and news-based economic uncertainty. VIX is instead chosen to be kept inside the model at the fourth selection step just after default type, seniority and industrial sector.

Default rates, usually considered to date as “the most relevant” systematic factor, are not selected by the algorithm for $\mathcal{M}_{4++}^{\mu,\phi}$ but they are included in $\mathcal{M}_{5++}^{\mu,\phi}$ where they do not display any significance. They are also chosen as the twelfth variable for order of importance, after all the uncertainty measures which display significance in $\mathcal{M}_{5++}^{\mu,\phi}$.

5.2.2. The role of uncertainty and default rates for varying dispersion

Uncertainty measures also prove to have a major role in determining the dispersion of recovery rates distributions. Financial uncertainty and news-based economic policy uncertainty are selected in the precision submodel of $\mathcal{M}^{\mu,\phi}_{4++}$, with financial uncertainty being chosen at the third selection step after the default type and industrial sector variables, hence being the most important systematic determinant also for the precision submodel. Both financial uncertainty and the news-based indicator are strongly significant in $\mathcal{M}^{\mu,\phi}_{4++}$ and feature a positive and negative sign respectively.

VIX index and the news-based measure are selected for the precision submodel of $\mathcal{M}^{\mu,\phi}_{5++}$. Only VIX shows to be strongly significant and features a positive sign. As for the selection procedure however, they are not the first systematic determinant to be selected. In fact, default rates already enter at the fourth selection step after default event, industrial sector and maturity.

Default rates are selected in the precision submodel of both model $\mathcal{M}^{\mu,\phi}_{4++}$ and $\mathcal{M}^{\mu,\phi}_{5++}$. They are strongly significant in both models and feature a negative sign.

From these analyses, it hence appears that, given μ, increases in volatility-based measures of uncertainty predict an increase in the precision parameter, hence a decrease in the dispersion of the distributions generating recovery rates. The contrary is true for the news-based measures.

Moreover, although default rate cannot be considered as a major determinant of average recovery rates when uncertainty measures are included in the models, it proves to be a key driver of the dispersion of recovery rates distributions. Increasing default rates predict higher variability of recovery rates distributions.

We also remark that seniority of the defaulted bond, industrial sector of the bond issuer and default type are selected in the first iterations of the algorithm for both the mean and precision submodels, supporting the evidences described in Section 3.3 about the role of these variables in determining the shapes of the corresponding conditional distributions.

6. Conclusion

Credit losses on bond portfolios strongly depend on three core factors: exposure at default, default probability and loss given default. While the first two received a lot of attention in the literature, the complement of the latter – the recovery rate – remained disregarded in spite of its primary importance. Moreover, among the few models developed so far, there are common flaws relating to the identification of bonds' recovery rates determinants. In fact, although researchers attained consensus on the variables responsible for the cross-sectional variations in recovery rates, there is still need to shed light on the determinants of the *systematic* fluctuations. Whereas past literature pinpoints default rates as the most important systematic determinant of recovery rates, we bring evidence which revisits this paradigm.

In this paper, we analyze the determinants of bonds' recovery rates using statistical techniques naturally suited to address recovery rates data specificities such as bounded support, skewed distributions and heteroskedasticity. These methods allow us to draw original conclusions on the role of debt characteristics and proxies of the business cycle in determining the shapes of recovery rates *distributions*.

We find that *economic uncertainty* – a concept which, to the best of our knowledge, is considered in this stream of literature for the first time – plays a key role. In particular, we find that measures of time-varying uncertainty play a greater role than default rates in explaining the systematic variations in recovery rates distributions. This evidence is still valid even after controlling for other proxies of the business cycle such as GDP, industrial production returns, loans' delinquency rates, a recession indicator and the level of the market index SP500. Higher uncertainty – as measured by four different types of proxy – is always associated with a lower *mean* of the recovery rates distributions.

Volatility-based proxies of uncertainty play the most relevant role with respect to the other types of measures. In fact, the explanatory power of volatility-based measures in explaining the means of recovery rates distributions is only second to default type and bonds' seniority and they outperform default rates in all the models we considered.

By contrast, we find default rates to be among the most relevant systematic factors regarding the *dispersion* of recovery rates distributions: it is the first variable selected after the default event, industrial sector of the bonds' issuer and maturity, all bond-specific characteristics. Increasing default rates predict higher dispersion in the recovery rates distributions. We register significant dispersion effects also for uncertainty proxies, with measures referring to different types of uncertainty featuring opposite signs.

It hence appears that even if default rates can no longer be identified as the main systematic determinant of average recovery rates while proxies of economic uncertainty are considered, they are still important systematic factors for explaining the variability in recovery rates.

Our results can help improving stochastic models for recovery rates used by practitioners for the computation of credit risk measures. In the light of the worldwide introduction of regulatory provisions allowing financial institutions to self-determine expected and unexpected credit losses (an option often exclusive to systemically important banks), taking into account these findings can have beneficial consequences for the stability of both single institutions and the overall financial system. This work also raises regulators' attention on new risk dimensions that must be considered when assessing the robustness of banks' internal models and the related results in terms of credit provisions and capital requirements.

The link found between bonds' recovery rates determinants and uncertainty measures also lays the groundwork for an interesting set of questions. Is it possible to create security-specific or firm-specific proxies of uncertainty? Why default rates and uncertainty measures convey different informations for the dispersion in recovery rates distributions? How is this information interpretable in an asset pricing framework? And in general, can our findings be integrated in a pricing model for defaulted securities? We leave all these questions for future research.

Online supplementary material

Technical appendix: Details regarding the likelihood equations, score vectors and Fisher's information matrices for both the classical and the variable dispersion beta regression models. Theoretical tables displaying functional forms of different link functions for the mean and precision parameters.

Figures of recovery rates distributions: Sample figures of recovery rates conditional distributions according to seniority of the defaulted bond, industrial sector of the bonds' issuer and default type.

Results of $\mathcal{M}^{\mu}_{4+}$, $\mathcal{M}^{\mu}_{5+}$ and $\mathcal{M}^{\mu}_{8+}$: Complete table of results for models $\mathcal{M}^{\mu}_{4+}$, $\mathcal{M}^{\mu}_{5+}$ and $\mathcal{M}^{\mu}_{8+}$ as defined in the main text.

References

Acharya, V., Bharath, S., and Srinivasan, A. (2007), "Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries," *Journal of Financial Economics*, 85, pp. 787–821.

Alexopoulos, M., and Cohen, J. (2015), "The power of print: uncertainty shocks, markets, and the economy," *International Review of Economics & Finance*, 40, pp. 8–28.

Altman, E., Brady, B., Resti, A., and Sironi, A. (2005), "The link between default and recovery rates: theory, empirical evidence, and implications," *The Journal of Business*, 78, pp. 2203–2227.

Altman, E., and Kalotay, E. (2014), "Ultimate recovery mixtures," *Journal of Banking & Finance*, 40, pp. 116–129.

Altman, E., and Karlin, B. (2009), "The re-emergence of distressed exchanges in corporate restructurings," *Journal of Credit Risk*, 5, pp. 43–56.

Altman, E., and Kishore, V. (1996), "Almost everything you wanted to know about recoveries on defaulted bonds," *Financial Analysts Journal*, Nov.-Dec., pp. 57–64.

Bachmann, R., Elstner, S., and Sims, E. (2013), "Uncertainty and economic activity: evidence from business survey data," *American Economic Journal: Macroeconomics*, 5, pp. 217–249.

Baker, S., Bloom, N., and Davis, S. (2015), "Measuring economic policy uncertainty," *NBER Working Paper Series*, 21633, pp. 1–75.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013), "Risk, uncertainty and monetary policy," *Journal of Monetary Economics*, 60, pp. 771–788.

Bernanke, B. S. (1983), "Irreversibility, uncertainty, and cyclical investment," *The Quarterly Journal of Economics*, 98, pp. 85–106.

Bertola, G., and Caballero, R. (1994), "Irreversibility and aggregate investment," *The Review of Economic Studies*, 61, pp. 223–246.

Bickel, P., and Doksum, K. (1981), "An analysis of transformations revisited," *Journal of the American Statistical Association*, 76, pp. 296–311.

Bloom, N. (2009), "The impact of uncertainty shocks," *Econometrica*, 77, pp. 623–685.

Bloom, N., Bond, S., and van Reenen, J. (2007), "Uncertainty and investment dynamics," *The Review of Economic Studies*, 74, pp. 391–415.

Boudreault, M., Gauthier, G., and Thomassin, T. (2013), "Recovery rate risk and credit spreads in a hybrid credit risk model," *Journal of Credit Risk*, 9, pp. 3–39.

Box, G. E. P., and Cox, D. (1964), "An analysis of transformations," *Journal of the Royal Statistical Society, Series B*, 26, pp. 211–252.

Bris, A., Welch, I., and Zhu, N. (2006), "The costs of bankruptcy: chapter 7 liquidation versus chapter 11 reorganization," *The Journal of Finance*, 61, pp. 1253–1303.

Bruche, M., and Gonzlez-Aguado, C. (2010), "Recovery rates, default probabilities and the credit cycle," *Journal of Banking & Finance*, 34, pp. 754–764.

Chulià, H., Guillén, M., and Uribe, J. M. (2017), "Measuring uncertainty in the stock market," *International Review of Economics & Finance*, 48, pp. 18–33.

Dixit, A. K., and Pindyck, R. S. (1994), *Investment Under Uncertainty*, Princeton: Princeton University Press.

Draper, N. R., and Cox, D. R. (1969), "On distributions and their transformation to normality," *Journal of the Royal Statistical Society, Series B*, 31, pp. 472–476.

ECB (2009), "Uncertainty and the economic prospects for the euro area," *ECB Economic Bulletin*, pp. 58–61.

——— (2016), "The impact of uncertainty on activity in the euro area," *ECB Economic Bulletin*, pp. 55–74.

Ferrari, S., and Cribari-Neto, F. (2004), "Beta regression for modelling rates and proportions," *Journal of Applied Statistics*, 31, pp. 799–815.

Franks, J. R., and Torous, W. N. (1994), "A comparison of financial recontracting in distressed exchanges and Chapter 11 reorganizations," *Journal of Financial Economics*, 35, pp. 349 – 370.

Frye, J. (2000), "Collateral damage," *Risk Magazine*, 13, pp. 91–94.

Gieseck, A., and Largent, Y. (2016), "The impact of macroeconomic uncertainty on activity in the euro area," *Review of Economics*, 67, pp. 25–52.

Girardi, A., and Reuter, A. (2017), "New uncertainty measures for the euro area using survey data," *Oxford Economic Papers*, 69, pp. 278–300.

Hu, Y.-T., and Perraudin, W. (2002), "The dependence of recovery rates and defaults," Tech. rep.

Jankowitsch, R., Nagler, F., and Subrahmanyam, M. G. (2014), "The determinants of recovery rates in the US corporate bond market," *Journal of Financial Economics*, 114, pp. 155 – 177.

John, J. A., and Draper, N. R. (1980), "An alternative family of transformations," *Journal of the Royal Statistical Society, Series C*, 29, pp. 190–197.

Jurado, K., Ludvigson, S., and Ng, S. (2015), "Measuring uncertainty," *American Economic Review*, 105, p. 11771216.

Knight, F. H. (1921), *Risk, uncertainty, and profit.*, Houghton Mifflin Company, MA: Hart, Schaffner & Marx.

Kose, M. A., and Terrones, M. (2012), "How does uncertainty affect economic performance?" *IMF World Economic Outlook*, pp. 49–53.

Leahy, J. V., and Whited, T. (1996), "The effect of uncertainty on investment: some stylized facts," *Journal of Money, Credit and Banking*, 28, pp. 64–83.

Ludvigson, S. C., Ma, S., Ng, S., and Olea, L. M. (2017), "Uncertainty and business cycles: exogenous impulse or endogenous response?" Working Paper.

Manly, B. F. (1976), "Exponential data transformation," *The Statistician*, 25, pp. 37–42.

McDonald, R. L., and Siegel, D. (1986), "The value of waiting to invest," *The Quarterly Journal of Economics*, 101, pp. 707–727.

Mora, N. (2012), "What determines creditor recovery rates?" *Federal Reserve Bank of Kansas City in its Journal Economic Review*, 79, pp. 109–123.

Nazemi, A., and Fabozzi, F. J. (2018), "Macroeconomic variable selection for creditor recovery rates," *Journal of Banking & Finance*, 89, pp. 14–25.

Ospina, R., and Ferrari, S. L. P. (2008), "Inflated beta distributions," *Statistical Papers*, 51, pp. 111–126.

Qi, M., and Zhao, X. (2011), "Comparison of modeling methods for loss given default," *Journal of Banking & Finance*, 35, pp. 2842–2855.

Rigby, R. A., and Stasinopoulos, D. (2008), "Generalized additive models for location scale and shape (GAMLSS) in R," *Journal of Statistical Software*, 23, pp. 1–46.

Rigby, R. A., and Stasinopoulos, D. M. (2005), "Generalized additive models for location, scale and shape," *Journal of the Royal Statistical Society, Series C*, 54, pp. 507–554.

Schuermann, T. (2004), "What do we know about loss given default?" *Wharton Financial Institutions Center Working Paper*, Working Paper No. 04-01.

Shleifer, A., and Vishny, R. (1992), "Liquidation values and debt capacity: a market equilibrium approach," *The Journal of Finance*, 47, pp. 1343–1366.

Simas, A., Barreto-Souza, W., and Rocha, A. (2010), "Improved estimators for a general class of beta regression models," *Computational Statistics & Data Analysis*, 54, pp. 348–366.

Smithson, M., and Verkuilen, J. (2006), "A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables," *Psychological Methods*, 11, pp. 54–71.

Soojin, J., and Rodrigo, S. (2017), "Macroeconomic uncertainty through the lens of professional forecasters," *Journal of Business & Economic Statistics*, pp. 1–11, Published Online at: https://doi.org/10.1080/07350015.2017.1356729.

Yao, X., Crook, J., and Andreeva, G. (2015), "Support vector regression for loss given default modelling," *European Journal of Operational Research*, 240, pp. 528–538.

Yeo, I., and Johnson, R. (2000), "A new family of power transformations to improve normality or symmetry," *Biometrika*, 87, pp. 954–959.

Zarnowitz, V., and Lambros, L. (1987), "Consensus and uncertainty in economic prediction," *Journal of Political Economy*, 95, pp. 591–621.

Tables

Table 1: Lower triangle of the correlation matrix of uncertainty measures.

	VIX	Fin. unc.	News-based EPU	CPI unc.	F.S.L. exp. unc.
VIX	100.00				
Fin. unc.	83.02	100.00			
News-based EPU	36.18	22.39	100.00		
CPI unc.	39.01	43.26	-3.73	100.00	
F./S./L. exp. unc.	-1.05	-9.33	25.15	32.13	100.00

NOTE: All displayed values are expressed in percentage.

Table 2: Summary statistics of our recovery rates sample.

Statistic	N	Max	Mean	Median	Min	Pctl(25)	Pctl(75)	St. Dev.
Recovery rate	1,831	100.00	30.28	21.00	0.01	10.00	47.50	25.60

NOTE: All displayed values except the sample size are expressed in percentage.

Table 3: Summary statistics of recovery rates according to the seniority of the defaulted bond.

	N	Min.	Median	Mean	Max.	St.Dev.	Skewness
Senior Secured	151	0.75	51.50	49.20	99.50	27.92	0.0485
Senior Unsecured	1180	0.30	19.50	29.66	99.50	25.49	0.9622
Senior Subordinated	413	0.01	20.75	26.70	100.00	22.72	1.0455
Subordinated	78	0.25	19.42	24.44	95.00	22.16	1.2025
Junior Subordinated	9	1.00	9.00	10.22	20.75	6.15	0.1511

NOTE: All displayed values except the sample size and skewness are expressed in percentage.

Table 4: Summary statistics of recovery rates according to the industrial sector of the bond's issuer.

	N	Min.	Median	Mean	Max.	St.Dev.	Skewness
Banking	45	0.25	13.50	17.81	57.00	19.46	0.9789
Capital Industries	420	0.13	24.62	31.42	99.25	24.84	0.7639
Consumer Industries	266	0.01	26.50	33.86	100.00	24.55	0.7426
Energy & Environment	112	1.00	35.00	35.61	90.50	21.27	0.5535
FIRE	334	0.13	10.00	22.74	94.50	26.68	1.6836
Media & Publishing	159	0.01	37.00	38.95	96.00	27.43	0.1417
Retail & Distribution	143	0.50	28.00	31.77	99.50	23.18	0.9018
Technology	212	0.38	12.75	18.27	99.50	17.28	2.1323
Transportation	103	1.75	23.00	33.36	95.25	23.37	0.9780
Utilities	37	13.99	83.00	76.01	91.00	16.07	-1.9797

NOTE: All displayed values except the sample size and skewness are expressed in percentage.

Table 5: Summary statistics of recovery rates according to the default type.

	N	Min.	Median	Mean	Max.	St.Dev.	Skewness
Bankruptcy	28	1.00	25.25	32.33	65.00	24.81	0.0833
Chapter 11	712	0.01	10.88	22.85	99.25	22.02	1.5121
Chapter 7	8	0.53	2.88	9.54	46.00	15.33	1.9541
Liquidated	7	13.00	15.38	15.12	16.57	1.35	-0.6822
Missed interest payment	796	0.01	26.00	32.84	100.00	23.88	0.7778
Missed principal & interest payments	66	8.25	65.50	57.68	91.00	29.22	-0.6340
Missed principal payment	25	1.00	39.00	44.40	96.53	26.60	0.3306
Others	5	3.00	9.00	16.33	51.63	19.92	1.4294
Payment moratorium	35	82.05	83.17	83.21	85.13	0.55	1.7748
Prepackaged Chapter 11	101	0.50	15.00	29.95	99.50	28.98	0.7595
Receivership	11	0.25	0.50	5.45	26.50	10.42	1.6362
Suspension of payments	37	5.00	18.50	25.84	99.00	19.41	1.6304

NOTE: All displayed values except the sample size and skewness are expressed in percentage.

Table 6: Beta regression models with one systematic determinant.

	Recovery rate distributions - Mean model							
	No systematic determinant		One systematic determinant					
	$\mathcal{M}_1^\mu$	$\mathcal{M}_2^\mu$	$\mathcal{M}_3^\mu$	$\mathcal{M}_4^\mu$	$\mathcal{M}_5^\mu$	$\mathcal{M}_6^\mu$	$\mathcal{M}_7^\mu$	$\mathcal{M}_8^\mu$
Constant	−2.425***	−2.130***	−2.255***	−1.884***	−2.099***	−1.990***	−2.132***	−2.302***
	(0.354)	(0.361)	(0.358)	(0.349)	(0.353)	(0.360)	(0.361)	(0.355)
Coupon		0.105***	0.112***	0.042	0.067**	0.101***	0.107***	0.104***
		(0.028)	(0.028)	(0.027)	(0.027)	(0.028)	(0.028)	(0.028)
Maturity		−0.083***	−0.067**	−0.072***	−0.090***	−0.073***	−0.081***	−0.097***
		(0.026)	(0.026)	(0.025)	(0.026)	(0.026)	(0.027)	(0.026)
Amr. def. rate			−0.202***					
			(0.024)					
Fin. unc.				−0.447***				
				(0.025)				
VIX					−0.346***			
					(0.024)			
CPI disagreement						−0.155***		
						(0.026)		
F./S./L. exp. disagr.							−0.025	
							(0.026)	
News-based EPU								−0.239***
								(0.024)
Observations	1,830	1,830	1,830	1,830	1,830	1,830	1,830	1,830
Seniority dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	**0.271**	**0.284**	**0.305**	$\boxed{\textbf{0.390}}$	**0.347**	**0.298**	**0.284**	**0.324**
Log Likelihood	697.644	711.089	743.566	859.771	800.580	728.823	711.525	760.875

NOTE: *p<0.1; **p<0.05; ***p<0.01. Logit link function is chosen for the mean model. All models include categorical variables for controlling for bonds' seniority and presence of backing, industrial sector of the bond issuer and default type. The reference categories are set to senior unsecured bonds without backing, financial insurance and real-estate companies and chapter 11 defaults. Model $\mathcal{M}_1^\mu$ only includes the constant. Model $\mathcal{M}_2^\mu$ also includes coupon level and maturity. Models $\mathcal{M}_3^\mu$ up to $\mathcal{M}_8^\mu$ all have the same number of regressors and additionally include one systematic determinant.

Table 7: Augmented beta regression models.

	Recovery rate distributions - Mean model			
	$\mathcal{M}^{\mu}_{2+}$	$\mathcal{M}^{\mu}_{3+}$	$\mathcal{M}^{\mu}_{4++}$	$\mathcal{M}^{\mu}_{5++}$
Constant	-1.877^{***}	-1.927^{***}	-1.924^{***}	-2.049^{***}
	(0.354)	(0.354)	(0.348)	(0.349)
Coupon	0.069^{**}	0.068^{**}	0.049^{*}	0.066^{**}
	(0.028)	(0.028)	(0.028)	(0.028)
Maturity	-0.069^{***}	-0.068^{***}	-0.085^{***}	-0.092^{***}
	(0.026)	(0.026)	(0.025)	(0.026)
Rec. indicator	-0.382^{***}	-0.323^{***}	-0.072	-0.147^{*}
	(0.075)	(0.080)	(0.087)	(0.087)
GDP	-0.017	-0.029	0.117^{***}	0.172^{***}
	(0.038)	(0.038)	(0.045)	(0.046)
IP return	0.164^{***}	0.158^{***}	0.088^{**}	0.134^{***}
	(0.036)	(0.036)	(0.036)	(0.036)
SP500	-0.246^{***}	-0.216^{***}	-0.139^{***}	-0.325^{***}
	(0.040)	(0.043)	(0.049)	(0.048)
Del. rate C./I. loans	-0.198^{***}	-0.143^{***}	0.053	-0.035
	(0.041)	(0.048)	(0.052)	(0.051)
Amr. def. rate		-0.075^{**}	-0.012	-0.044
		(0.033)	(0.033)	(0.033)
Fin. unc.			-0.338^{***}	
			(0.034)	
VIX				-0.233^{***}
				(0.030)
CPI disagreement			-0.028	-0.051
			(0.034)	(0.034)
F./S./L. exp. disagr.			-0.044	-0.074^{*}
			(0.040)	(0.041)
News-based EPU			-0.132^{***}	-0.117^{***}
			(0.028)	(0.029)
Observations	1,830	1,830	1,830	1,830
Seniority dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Event dummies	Yes	Yes	Yes	Yes
Pseudo-R^2	**0.352**	**0.353**	$\boxed{\mathbf{0.404}}$	**0.389**
Log Likelihood	802.542	805.040	881.247	862.116

NOTE: *p<0.1; **p<0.05; ***p<0.01. Logit link function is chosen for the mean model. All models includes categorical variables for controlling for bonds' seniority and presence of backing, industrial sector of the bond issuer and default type. The reference categories are set to senior unsecured bonds without backing, financial insurance and real-estate companies and chapter 11 defaults. Model $\mathcal{M}^{\mu}_{2+}$ only includes a recession indicator, a monthly polynomial interpolation of the GDP for United States, monthly IP returns, level of the SP500 and a monthly polynomial interpolation of delinquency rates for commercial and industrial loans. Model $\mathcal{M}^{\mu}_{3+}$ additionally includes the monthly American default rate. Model $\mathcal{M}^{\mu}_{4++}$ and $\mathcal{M}^{\mu}_{5++}$ have the same dimensionality and they include different types of proxies of uncertainty. Due to high correlation between the financial uncertainty indicator and the VIX index, Model $\mathcal{M}^{\mu}_{4++}$ and $\mathcal{M}^{\mu}_{5++}$ only include one of the two variables.

Table 8: Variable dispersion beta regression models.

	Recovery rate distributions (Mean submodel)	
	$\mathcal{M}^{\mu,\phi}_{4++}$	$\mathcal{M}^{\mu,\phi}_{5++}$
Constant	−3.486***	−3.465***
	(0.165)	(0.145)
Coupon	0.057***	0.077***
	(0.019)	(0.021)
Maturity	−0.082***	−0.101***
	(0.014)	(0.015)
Rec. indicator		−0.287***
		(0.086)
GDP	−0.007	0.141***
	(0.037)	(0.044)
IP return	0.107***	0.133***
	(0.029)	(0.035)
SP500	−0.023	−0.284***
	(0.031)	(0.044)
Del. rates C./I. loans		
Amr. def. rate		−0.003
		(0.032)
Fin. unc.	−0.458***	
	(0.029)	
VIX		−0.241***
		(0.029)
News-based EPU	−0.058**	−0.116***
	(0.026)	(0.028)
F./S./L. exp. disagr.		−0.086**
		(0.038)
CPI disagreement		−0.046
		(0.034)
Observations	1,830	1,830
Seniority dummies	Yes	Yes
Industry dummies	Yes	Yes
Event dummies	Yes	Yes
Pseudo-R^2	0.376	0.370
Log Likelihood	1,217.484	1,147.315

	Recovery rate distributions (Precision submodel)	
	$\mathcal{M}^{\mu,\phi}_{4++}$	$\mathcal{M}^{\mu,\phi}_{5++}$
Constant	3.718***	4.189***
	(0.500)	(0.507)
Coupon	−0.087**	−0.075*
	(0.039)	(0.039)
Maturity	0.325***	0.371***
	(0.037)	(0.036)
Rec. indicator		0.266**
		(0.115)
GDP	0.120**	−0.032
	(0.056)	(0.043)
IP return	−0.123***	−0.071
	(0.045)	(0.051)
SP500	−0.116*	
	(0.061)	
Del. rates C./I. loans	0.282***	
	(0.068)	
Amr. def. rate	−0.420***	−0.215***
	(0.043)	(0.040)
Fin. unc.	0.463***	
	(0.046)	
VIX		0.213***
		(0.041)
News-based EPU	−0.211***	−0.046
	(0.038)	(0.038)
F./S./L. exp. disagr.		
CPI disagreement		
Seniority dummies	Yes	Yes
Industry dummies	Yes	Yes
Event dummies	Yes	Yes

NOTE: Models feature a logit link for the mean submodel and log-link for the precision submodel. The two models are compared across the mean submodel and the precision submodel. Variables have been selected with backward step-GAIC. The starting models for the variable selection procedure generating $\mathcal{M}^{\mu,\phi}_{4++}$ and $\mathcal{M}^{\mu,\phi}_{5++}$ are respectively represented by Model $\mathcal{M}^{\mu}_{4++}$ and $\mathcal{M}^{\mu}_{5++}$ of Table 7. Note: *p<0.1; **p<0.05; ***p<0.01.