econometrics

Article

Exploring Industry-Distress Effects on Loan Recovery: A Double Machine Learning Approach for Quantiles

Hui-Ching Chuang [1] and Jau-er Chen [2,3,*]

1 College of Management, Yuan Ze University, Taoyuan 320, Taiwan
2 Department of Economics, Senshu University, Kawasaki 214-8580, Kanagawa, Japan
3 Center for Research in Econometric Theory and Applications, National Taiwan University, Taipei 10617, Taiwan
* Correspondence: jauer@isc.senshu-u.ac.jp; Tel.: +81-44-900-7983

Abstract: In this study, we explore the effect of industry distress on recovery rates by using the unconditional quantile regression (UQR). The UQR provides better interpretative and thus policy-relevant information on the predictive effect of the target variable than the conditional quantile regression. To deal with a broad set of macroeconomic and industry variables, we use the lasso-based double selection to estimate the predictive effects of industry distress and select relevant variables. Our sample consists of 5334 debt and loan instruments in Moody's Default and Recovery Database from 1990 to 2017. The results show that industry distress decreases recovery rates from 15.80% to 2.94% for the 15th to 55th percentile range and slightly increases the recovery rates in the lower and the upper tails. The UQR provide quantitative measurements to the loss given default during a downturn that the Basel Capital Accord requires.

Keywords: loss given default; recentered influence function; quantile regression; double machine learning; lasso

Citation: Chuang, Hui-Ching, and Jau-er Chen. 2023. Exploring Industry-Distress Effects on Loan Recovery: A Double Machine Learning Approach for Quantiles. *Econometrics* 11: 6. https://doi.org/10.3390/econometrics11010006

Academic Editor: Martin Biewen

Received: 7 November 2022
Revised: 5 February 2023
Accepted: 10 February 2023
Published: 14 February 2023

1. Introduction

Default recovery rate (R hereafter), or loss given default (LGD, one minus R), is one of the critical components in credit risk management. Basel II requires financial institutions adopting the advanced internal rating based approach to provide adequate R/LGD estimate, cf. Basel Committee on Banking Supervision, BCBS (2004) and BCBS (2005), and according to Articles 181(b) and 182(b) of the EU Capital Requirements Regulation No 575/2013, institutions need to use R/LGD estimates on an economic downturn if those are more conservative than the long-run averages.[1] In order to quantify the downturn LGD, the European Banking Authority publishes the guidelines on the estimation of the downturn LGD, cf. EBA (2018) and EBA (2019). How to measure the effect of the economic downturn on LGD is, therefore, an essential task for both industry practitioners and academic researchers.

The conditional quantile regression (CQR) of Koenker and Bassett (1978) has been used to model the downturn default recovery; see, e.g., Somers and Whittaker (2007) and Krüger and Rösch (2017). However, CQR models the conditional quantile function of recovery rate, and that conditional distribution depends on the researcher's specified covariates. Thus, the interpretation of the CQR coefficients is varying when the model has different covariates. For example, if we want to compare the distress effects on two models with different covariates, these two coefficients of industry distress are not comparable, because the quantiles we refer to are of two different conditional distributions. Furthermore, suppose that we estimate a CQR model at the 50th quantile with industry distress and type of debt instrument dummies. The distress coefficient is not the effect of distress at the 50th quantile of the recovery rate distribution. Rather, it is the average of the effects at the 50th quantiles of the distributions for all types of instruments.

In this study, we estimate the effect of industry distress on recovery rates by using the unconditional quantile regression (UQR) proposed in Firpo et al. (2009). The UQR provides better interpretative and thus policy-relevant information on the predictive effect of the covariates than the conditional quantile regression. The UQR measures the effect at specific quantiles of the unconditional recovery rate but not that within subcategories. Furthermore, different model specifications only reflect the model selection and do not affect the interpretation of the coefficients. Borah and Basu (2013), Maclean et al. (2014), and Porter (2015) provide more comparison studies between CQR and UQR in economics studies.

We study the effect of industry distress on recovery rates by using 5334 debt and loan instruments from Moody's Default and Recovery Database for the period from 1990 to 2017. Following Acharya et al. (2007), we approximate the issuers' industry distress if the median stock returns of the firms in the same industry are less than -30%. We investigate the effects of the following debt characteristics: collateral status, collateral quality, instrument type, and debt cushion level. To control for macroeconomic and industry conditions, we collect 72 industry variables and 130 macroeconomic variables. The variables we choose to proxy for macroeconomic conditions are similar to the ones selected in Nazemi and Fabozzi (2018) and Krüger and Rösch (2017). The selection of the macroeconomic and industry variables in the UQR is vital to the precision of the estimates of industry distress. These variables simultaneously affect the recovery rate (outcome variable) and the industry distress (target variable). To refine our selection methodology, we adapt the lasso-based double selection procedure of Belloni et al. (2014a) for conditional mean models and Chen et al. (2021) for conditional quantile models to the UQR. By implementing this procedure, we select 12 variables that are common across all recovery rate quantiles with a maximum of 22 selected variables for the 25th quantile model. The relatively small number of the variables means that the proposed selection method benefits from model shrinkage to a large extent.

Our results show that industry distress decreases recovery rates significantly from 15.80% to 2.94% from the 15th to the 55th percentile range and slightly increases the recovery rates in the lower and upper tails. In contrast to the estimates from a CQR with the same covariates, the industry distress shows no significant effect in almost all conditional quantiles of the recovery rate. We show that the status of the collateral and the level of the debt cushion have heterogeneous effects on different quantiles of recovery rates, with the most significant being between the 50th and 55th percentiles, respectively.

Regarding the estimation of unconditional quantile effects in the literature, our paper is related to Sasaki et al. (2022). They found that the unconditional quantile effects can be represented as the average derivative estimator and employ a semiparametric influence adjustment term to correct for nonparametric estimation errors from the regularized preliminary estimation in the presence of high-dimensional covariates. To make the implementation more practical and accessible for empirical researchers, we adopt the recentered influence function approach and model the RIF as a linear function of covariates when estimating unconditional quantile effects through our double selection procedure.

Our study contributes to the literature in several ways. First, we add to the literature by exploring the heterogeneous effect of industry risk on the recovery rate. Shleifer and Vishny (1992) identify that the fire sale channel is a critical factor that affects the firm's asset liquidation and recovery value. When a firm experiences financial distress, its industry peers may also suffer, leading to discounts on assets due to market illiquidity. Empirical evidence of the relationship between industry risk and recovery rates can be found in works by Acharya et al. (2007), James and Kizilaslan (2014), and Chang et al. (2020). Gambetti et al. (2019) also address the effect of economic uncertainty. Our study estimates the impact of industry distress at each quantile of the unconditional recovery rates, offering a more comprehensive picture of the fire sale effect and the effect of economic downturn on asset liquidation values.

Second, we complement the studies of Somers and Whittaker (2007), Siao et al. (2016), and Krüger and Rösch (2017), who use a conditional quantile regression or the modified logistic quantile regression to model the loss given default (LGD) values during a downturn. Our study provides additional measurements for the effect of industry distress on the unconditional quantiles of recovery rates. These measurements are interpretative and thus policy-relevant when adjusting the LGD values from normal values during downturns. Furthermore, our study focuses on exploring heterogeneous effects across quantiles in a data-rich environment with high-dimensional covariates. We also add to the existing literature on LGD studies by exploring the estimation target beyond the conditional mean level; see Bastos (2010), Qi and Zhao (2011), Gürtler and Hibbeln (2013), Hartmann-Wendels et al. (2014), and others.

Third, we employ a lasso-based double selection procedure to select macroeconomic and industry variables. This approach is based on the work of Belloni et al. (2014a), Belloni et al. (2014b) for conditional mean models, and Chen and Hsiang (2019), as well as Chen et al. (2021) for conditional quantile models. It has also been used in asset pricing applications, as seen in Feng et al. (2020). Our double selection procedure improves upon the single selection used by Nazemi and Fabozzi (2018) by accommodating the scenario of a large number of characteristics and enabling examination of their effect through a smaller set of superior variables. This approach preserves the original meanings of the candidate variables, unlike factor analysis or principal component analysis, c.f. Nazemi et al. (2018). Recent research has also applied advanced machine learning and deep learning to investigate the factors determining recovery rates, such as in the studies by Kellner et al. (2022) and Nazemi et al. (2022).

This paper proceeds as follows: Section 2 introduces the unconditional and conditional quantile regressions and the variable selection procedure. Section 3 describes the descriptive statistics for the recovery data and estimation results. Section 4 concludes the paper. S1 lists the recovery rate by the Fama–French 30 industry classification. S2 lists the definitions of the industry-specific variables.[2]

2. Unconditional versus Conditional Quantile Regression

In this section, we introduce the unconditional quantile regression of Firpo et al. (2009) and review its difference from the conditional quantile regression of Koenker (2005). We also describe the lasso-based double selection approach.

2.1. Unconditional Quantile Regression

Let R denote the outcome of interest and recovery rates, and let $F_R(r)$ be the distribution function of R. We are interested in the effect of covariate X on R. If X is a binary variable, then the (unconditional) distribution can be written as the weighted sum of the conditional distributions:

$$F_R(r) = p\, F_{R|X}(r|X=1) + (1-p)\, F_{R|X}(r|X=0).$$

where $p = Pr(X=1)$, and $F_{R|X}$ is the conditional distribution of R on X. For example, if we want to know the effect of industry distress on the recovery rate, then X is the distress dummy that equals one if the industry is in the state of distress, and zero otherwise. An ordinary least square (OLS) estimates the predictive effect of X on the conditional mean of R, i.e., $\beta_{OLS} = E_{R|X}(r|X=1) - E_{R|X}(r|X=0)$. We know that

$$E(R) = p\, E_{R|X}(r|X=1) + (1-p)\, E_{R|X}(r|X=0);$$
$$\frac{dE(R)}{dp} = E_{R|X}(r|X=1) - E_{R|X}(r|X=0).$$

Thus, an OLS coefficient can also reflect the predictive effect of X on the unconditional mean of R, i.e., $\frac{dE(R)}{dp}$.

In addition to the mean level, we are interested in the effect of X on the quantiles of R. The CQR estimate is defined as

$$\beta_\tau^{CQR} := Q_\tau(R|X=1) - Q_\tau(R|X=0),$$

where $Q_\tau(R|X) = F_{R|X}^{-1}(\tau|X) := \inf\left\{z : F_{R|X}(z|X) \geq \tau\right\}$ is the τth conditional quantile of R on X. Let the τth (unconditional) quantile of R be $q_\tau := F_R^{-1}(\tau)$ and the effect of X to the τth quantile of R be dq_τ/dp. Unlike the OLS coefficient, the CQR estimate has no such duality interpretation as an OLS coefficient, that is, β_τ^{CQR} is generally not equal to dq_τ/dp. Firpo et al. (2009) show that the unconditional quantile partial effect (UQPE) is of the form:

$$dq_\tau/dp = [Pr(R > q_\tau|X=1) - Pr(R > q_\tau|X=0)]/f_R(q_\tau),$$

and further expresses the UQPE as the weighted average of conditional quantile partial effect.

The UQR is based on the recentered influence function (RIF) of the dependent variable.

$$RIF(R, q_{\tau_0}) = q_{\tau_0} + \frac{q_{\tau_0} + (\tau_0 - \mathbf{1}_{\{R \leq q_{\tau_0}\}})}{f_R(q_{\tau_0})}. \tag{1}$$

where q_{τ_0} is the τ_0th quantile value of the dependent variable, and $\mathbf{1}_{\{A\}}$ is the indication function that equals one if event A is true, and zero otherwise. $f_R(q_{\tau_0})$ is the density of R evaluated at q_{τ_0}. In practice, we use a kernel density estimator (KDE), $\hat{f}_R(q_{\tau_0})$, to replace $f_R(q_{\tau_0})$. For instance, the KDE with Gaussian kernel and bandwidth h is in the form of

$$\hat{f}(x) = \frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x - x_j}{h}\right).$$

and $k(z) = \frac{1}{\sqrt{2\pi}} \exp \frac{-z^2}{2}$ is the Gaussian density.[3] The UQR estimator is defined as the coefficient vector from the linear regression of the RIF of R on the covariate X.

$$RIF(R, q_{\tau_0}) = X\beta_{\tau_0}^{UQR} + \epsilon.$$

The coefficient $\beta_{\tau_0}^{UQR}$ corresponds to the target of interest $dq_\tau/dp|_{\tau=\tau_0}$ in the linear specification. We use the bootstrapping method to obtain the standard error of the UQR estimators.

Compared with the UQR, the CQR proposed in Koenker and Bassett (1978) estimates the conditional quantile function. The linear setting is

$$Q_\tau(R|X) = X\beta_\tau^{CQR}.$$

The unknown parameter β_τ^{CQR} can be estimated by minimizing the following objective function:

$$\sum_{i=1}^{n} \rho_\tau\left(R_i - X_i\,\beta_\tau^{CQR}\right),$$

with

$$\rho_\tau(z) = \begin{cases} \tau z & \text{if } z \geq 0; \\ (1-\tau)|z| & \text{otherwise.} \end{cases}$$

The function $\rho_\tau(\cdot)$ is called the check function or the tilted absolute value function. Unlike the OLS with equal weights on the residuals, this objective function gives different weights to the residuals. The linear programming method is applied to solve this minimization problem. Koenker (2005) shows that for $\tau \in (0,1)$ under mild regularity conditions, the CQR estimator $\hat{\beta}_\tau^{CQR}$ satisfies

$$\sqrt{n}\left(\hat{\beta}_{\tau}^{CQR} - \beta_{\tau}^{CQR}\right) \longrightarrow \mathcal{N}\left(0, \tau(1-\tau)D^{-1}\,\Omega_X\,D^{-1}\right)$$

with $D = E[f_R(X\beta_{\tau}^{CQR})XX']$ and $\Omega_X = E(X'X)$. One can estimate the standard errors of the estimator with the kernel method; however, to be consistent with the UQR estimators, we apply the bootstrapping method to obtain the standard errors of the CQR coefficient.

2.2. The Lasso-Based Double Selection Procedure

We adopt a double selection procedure by following the spirit of Belloni et al. (2014a) and Chen et al. (2021) to correct for the variable selection bias. We divide the explanatory variables into two categories but exclude the industry distress dummy: required variables and to-be-selected variables. The first variables are the main interests of researchers, such as debt and loan characteristics and maybe their interaction with industry distress. We denote them as $\mathbf{X}_{\text{must}}$. The second variables are less informative variables or the variables that are not the main interests of researchers, such as the macroeconomic and industry conditions, which are denoted as $\mathbf{X}_{\text{tobe}}$. Our proposed procedure aims to control for the effect of $\mathbf{X}_{\text{must}}$ on the recovery rate and to be as parsimonious as possible in selecting the to-be-selected variables.

The steps for the lasso-based double selection proceed are described as follows:

1. Preselection partialling-out:
   (a) Run an OLS with the RIF of the recovery rate in Equation (1) on $\mathbf{X}_{\text{must}}$ to obtain residuals ρ_{RIF}.
   (b) Run an OLS with the industry distress dummy on $\mathbf{X}_{\text{must}}$ to obtain residuals ρ_d.
   (c) For each variable j in the $\mathbf{X}_{\text{tobe}}$, run an OLS of $X_{\text{tobe},j}$ on $\mathbf{X}_{\text{must}}$ to obtain residuals $\rho_{\text{tobe},j}$. We denote $\boldsymbol{\rho}_{\text{tobe}}$ as the result matrix in this step.
2. Double selection:
   (a) Run a lasso regression on the ρ_{RIF} and $\boldsymbol{\rho}_{\text{tobe}}$. This step selects the to-be-selected variables that best explain the residuals of the RIF, ρ_{RIF}. As we already control for the effect of $\mathbf{X}_{\text{must}}$ in step 1(a) and 1(c), this step aims to select the $\mathbf{X}_{\text{tobe}}$ with the most predictive power for the reaming unexplained (RIF of the) recovery rates. Denote $\hat{I}_1$ as the set of indices corresponding to the selected variables in this step.
   (b) Run a lasso regression on the ρ_d and $\boldsymbol{\rho}_{\text{tobe}}$. This step selects the to-be-selected variables that best explain the residuals of industry distress, ρ_d. Because we already controlled for the effect of $\mathbf{X}_{\text{must}}$ in step 1(b) for the industry distress, this step aims to select the $\mathbf{X}_{\text{tobe}}$ with the most predictive power to the remaining unexplained industry distress. Denote $\hat{I}_2$ as the set of indices corresponding to the selected variables in this step.[4]
3. Postselection estimation: Run an OLS with the RIF of the recovery rate on the industry distress dummy, $\mathbf{X}_{\text{must}}$ and $\mathbf{X}_{\text{selected}}$, where $\mathbf{X}_{\text{selected}}$ is the subset of $\mathbf{X}_{\text{tobe}}$ with the variable indexed as the union of $\hat{I}_1$ and $\hat{I}_2$.

Following Belloni et al. (2014a) and Chen et al. (2021), we make the industry distress the target variable and the RIF of the recovery rate the outcome variable and select the variables with the most predictive power for both of them. The above-proposed procedure further accommodates the situation in which there is a relatively small set of superior variables ($\mathbf{X}_{\text{must}}$) that researchers would like to be included in the model while there is a rather large set of less informed characteristics ($\mathbf{X}_{\text{tobe}}$) to be controlled for. The preselection steps draw out the effects of $\mathbf{X}_{\text{must}}$ on the outcome, the target, and the $\mathbf{X}_{\text{tobe}}$ variables; the double selection steps find the union of variables with the best ability to predict the (remaining unexplained) outcome and the target variables.

3. Empirical Results

3.1. Recovery Data

We collect recovery data from Moody's Default and Recovery Database (DRD). Moody's DRD provides recovery information for instruments such as instrument type, default and settlement dates, collateral status, and industry classifications. Table 1 lists the definitions of the variables used in this study. There are three types of recovery rates in Moody's DRD: the discount liquidity, the discount settlement, and the trading price. The discount liquidity and settlement are the total nominal liquidity recovery and the settlement recovery amount that is discounted back from the trading date to the last cash paid date, with the defaulted instrument's effective interest rate as the discount rate. The trading price is the nominal recovery value that is discounted from the trading date to the instrument's last day that cash was paid using the effective interest rate of the predefaulted instrument. We use Moody's recommended discounted recovery rate of its Investor Service (MIS) that is based on the internal research standards. The total number of instruments in our study is 5334 from 1990 to 2017.

Table 1. Definitions of recovery rate and instrument characteristics.

Variable	Definitions
collateral	A dummy variable equals one if the debt has collateral, and zero otherwise.
industry	Instrument's issuer's industry is classified by the 30 Fama–French industry portfolio classification.
industry distress	A dummy variable that equals one if the median stock returns of the firms with the same industry classification to the instrument's issuer is less than −30% as Acharya et al. (2007). The year of the annual stock return is measured as the year at the middlepoint between default and emergence date of the instrument.
instrument type	Instrument type. One of Revolver, Term Loan, Senior Secured Bonds, Senior Subordinated Bonds, Senior Unsecured Bonds, Subordinated Bonds, Junior Subordinated Bonds.
percentage below	At the time of default, debt below is the total dollar amount outstanding of all defaulted debt that is contractually subordinate to the current instrument. Percentage below is debt below divided by the total issuer's debt.
rank	Collateral quality rank. Moody's DRD database ranks instrument's collateral quality as 1, 2, $\cdots$, 8. We define the rank as 1, 2, 3, 4 by winsorizing the original rank at 4 due to the limited observations with ranking above 4.
recovery rate	Moody's recommended recovery rate. Moody's Investor Service (MIS), based on internal research standards, recommend the recovery rate based on either trading price, liquidity, or settlement discounted recovery.
year	Instrument's year dummy is created as the year at the middlepoint between default and emergence date of the instrument.

Figure 1 plots the histogram and density estimate of the recovery rate. The rate has a bimodal shape and, of the observations, 5.62% and 35.51% are concentrated at zero recovery and full recovery. There are 1.52% of observations which are higher than the full recovery possible due to the additional fees required to the settlement payments and therefore exceeding the original debt amount. The median and mean of the recovery rate are 65.41% and 59.41%, respectively. The recovery rate reaches full recovery as early as the 63rd quantile. Table 2 lists the time trend of quartiles for recovery rates, and Figure 2 shows the time series pattern of recovery rate and the number of the defaulted instruments. We find the well-documented inverse relation between default count and (the median of) recovery rates through the years in Figure 2, cf. Altman et al. (2005), Bruche and González-Aguado (2010), Chava et al. (2011) and, Jankowitsch et al. (2014).

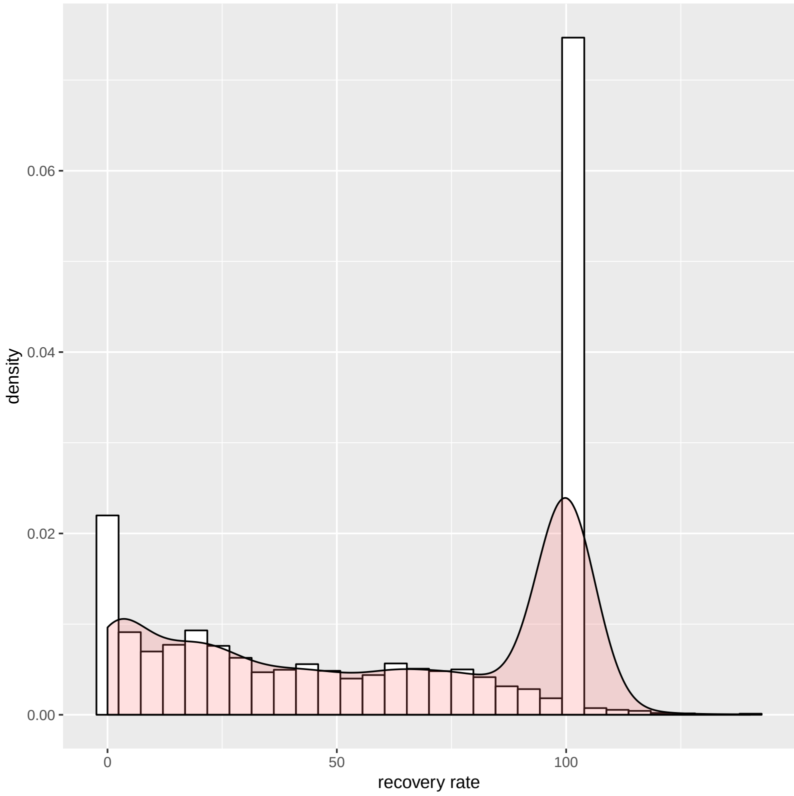

Figure 1. Histogram and density estimates of the recovery rate.

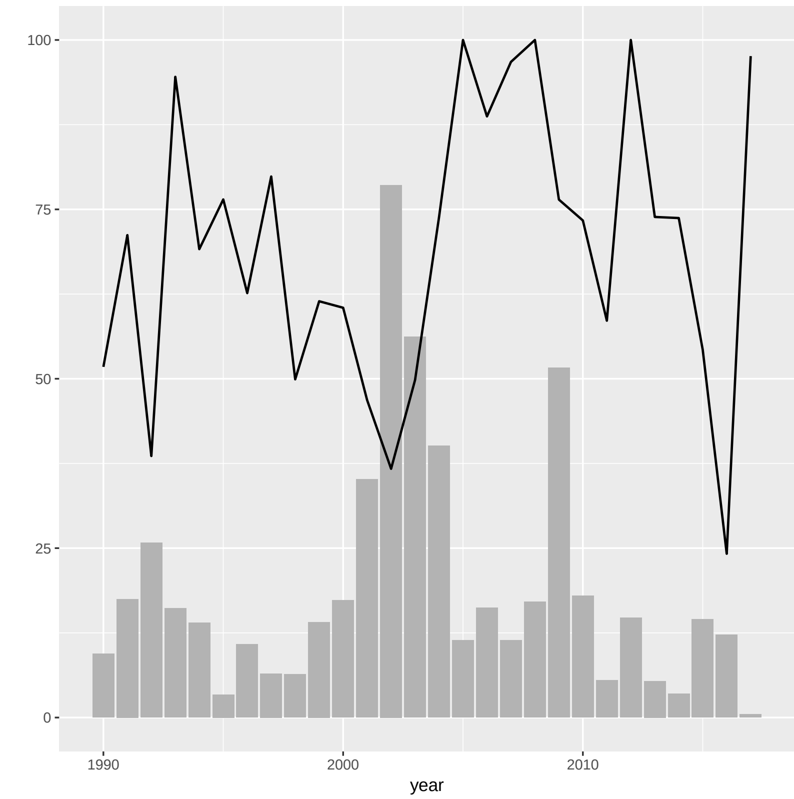

Figure 2. Time trend of median recovery rate and defaulted instrument count. Line is the median of the recovery rates in a year and bar is the defaulted instruments number divided by 10.

Table 2. Quantiles of recovery rate through years.

Year	10%	25%	50%	75%	Avg.	Obs.	Freq.
Recovery Rate	**1.93**	**20.61**	**65.41**	**100.00**	**59.41**	**5334**	**100.00**
1990	6.23	18.78	51.76	100.00	54.56	94	1.76
1991	1.90	16.00	71.20	100.00	59.66	175	3.28
1992	2.25	16.05	38.60	83.44	48.91	258	4.84
1993	2.51	23.42	94.56	100.00	66.24	161	3.02
1994	0.95	24.10	69.13	100.00	60.52	140	2.62
1995	0.15	30.64	76.47	100.00	64.23	34	0.64
1996	8.89	29.55	62.64	100.00	63.22	108	2.02
1997	6.26	16.18	79.85	100.00	63.97	65	1.22
1998	6.34	21.81	49.92	100.00	57.46	64	1.20
1999	5.19	20.70	61.44	100.00	59.88	141	2.64
2000	0.75	14.11	60.48	100.00	55.97	173	3.24
2001	0.47	7.48	46.86	100.00	51.23	352	6.60
2002	1.60	15.27	36.71	100.00	50.87	786	14.74
2003	1.32	21.83	49.82	100.00	53.09	562	10.54
2004	15.85	52.48	73.85	100.00	70.07	402	7.52
2005	24.82	66.12	100.00	100.00	79.23	114	2.14
2006	17.64	53.31	88.72	100.00	74.63	162	3.04
2007	3.67	56.64	96.77	100.00	76.97	114	2.14
2008	8.49	37.53	100.00	100.00	68.01	171	3.21
2009	1.15	21.06	76.44	100.00	62.60	516	9.67
2010	1.56	29.26	73.36	100.00	63.49	180	3.37
2011	0.00	0.44	58.57	100.00	53.19	55	1.03
2012	4.25	42.91	100.00	100.00	71.72	147	2.76
2013	1.37	39.58	73.88	100.00	63.50	54	1.01
2014	0.42	26.18	73.72	100.00	60.83	35	0.66
2015	11.06	24.53	54.26	100.00	60.01	145	2.72
2016	0.63	7.05	24.18	92.33	41.98	122	2.29
2017	71.36	71.64	97.62	100.00	88.09	5	0.09

Table 3 lists the summary statistics of the quantiles of the recovery rate by different subcategories. Panels A and B contain the recovery rates by the status and quality rank of the collateral, respectively. They signify that the collateralized instruments have higher recovery rates than the noncollateralized ones; the higher the quality ranking of the collateral, the more significant the recovery rate. Panel C shows the recovery rates by instrument types. The revolver has the highest mean of recovery rates, while the mean of the junior subordinated bond is the lowest. Panel D divides the sample into whether the year that the issuer goes into default is a year of industry distress or not. Industry distress is a dummy variable that equals one if the median stock return of the firms with the same industry classification as the instrument's issuer is less than -30%, as in Acharya et al. (2007). The year of the stock return is measured from the midpoint between the default date and the beginning of the instrument. We observe that the industry distress lowers the recovery rate by 5.39% on average. However, it slightly increases the recovery rate in the 10% quantile. S1 lists the recovery rate by the Fama–French 30 industry classification. The utility sector has the highest average recovery rates at 85.58%, while mining has the lowest average recovery rates at 50.41%.

The interaction between distress and instrument type is an intriguing issue that we observe in the data. Panel E of Table 3 shows that industry distress has a different effect on different types of instruments. We plot the density estimates of recovery rates of instrument types and the industry distress indicator in Figure 3. The figure shows that the distributions of recovery rates vary for the conditioning variables. On average, if the issuer's industry is in distress, the recovery rate is lower; however, its effect on different types of instruments is different, for example, for senior unsecured bonds, the industry distress lowers the average recovery rate by 18.46%. The density estimate shifts to the left if there is distress; however, the density estimate of the subordinated bonds shifts to the right when there is distress.

Table 3. Summary statistics of recovery rate.

Quantile	10%	25%	50%	75%	90%	Avg.	Obs.
Recovery Rate	**1.93**	**20.61**	**65.41**	**100.00**	**100.00**	**59.41**	**5334**
Panel A: Collateral status							
No	0.00	5.10	29.01	74.27	100.00	41.00	2582
Yes	20.88	53.26	100.00	100.00	100.00	76.67	2752
Panel B: Collateral quality rank							
1	24.14	58.17	100.00	100.00	100.00	78.07	2551
2	0.93	14.19	37.85	88.49	100.00	47.45	1791
3	0.00	1.44	18.10	60.73	100.00	34.05	659
4	0.00	0.82	16.31	64.20	98.52	30.89	333
Panel C: Instrument type							
Junior Subordinated	0.00	0.00	3.35	21.60	88.88	20.49	73
Revolver	40.00	80.72	100.00	100.00	100.00	86.29	1112
Senior Secured	18.44	23.72	62.24	100.00	100.00	61.95	706
Senior Subordinated	0.00	1.36	14.96	49.72	82.09	28.37	508
Senior Unsecured	1.22	12.52	39.71	88.92	100.00	47.86	1528
Subordinated	0.00	0.10	14.67	48.60	98.06	28.55	365
Term Loan	16.52	47.09	100.00	100.00	100.00	74.58	1042
Panel D: Industry distress							
No	1.96	20.91	66.87	100.00	100.00	60.22	4527
Yes	2.12	12.99	58.38	100.00	100.00	54.83	807
Panel E: Instrument type × Industry distress							
Junior Subordinated × Distress (No)	0.00	0.00	3.31	21.73	90.75	21.24	70
Junior Subordinated × Distress (Yes)	0.87	2.18	4.35	4.40	4.43	2.93	3
Revolver × Distress (No)	42.08	80.99	100.00	100.00	100.00	86.65	949
Revolver × Distress (Yes)	29.35	78.42	100.00	100.00	100.00	84.23	163
Senior Secured × Distress (No)	18.78	24.07	63.31	100.00	100.00	62.52	655
Senior Secured × Distress (Yes)	12.78	18.64	50.42	99.13	100.00	54.67	51
Senior Subordinated × Distress (No)	0.00	1.61	14.85	48.62	80.27	28.07	446
Senior Subordinated × Distress (Yes)	0.03	0.75	15.63	54.54	92.55	30.57	62
Senior Unsecured × Distress (No)	1.41	15.32	47.29	100.00	100.00	51.45	1231
Senior Unsecured × Distress (Yes)	0.43	8.26	24.36	61.69	77.92	32.98	297
Subordinated × Distress (No)	0.00	0.11	14.85	45.04	88.47	27.33	323
Subordinated × Distress (Yes)	0.00	0.00	8.26	80.35	100.00	37.97	42
Term Loan × Distress (No)	17.78	46.93	100.00	100.00	100.00	74.18	853
Term Loan × Distress (Yes)	16.33	48.64	100.00	100.00	100.00	76.38	189

3.2. Variable Selection

We select the macroeconomic variables based on the list in Nazemi and Fabozzi (2018), but we use the public website of the Federal Reserve Bank of St. Louis only. We also follow the selection of Krüger and Rösch (2017) and add the S&P500 return, TED spread, and the VIX to the list of macroeconomic variables.[5] After matching to the default month of each instrument and deleting insufficient observations, we have 130 macroeconomic variables. S2 lists the definitions of all macroeconomic variables used in our study. As industry variables, we consider 70 variables from Wharton Research Data Services Industry Financial Ratio Suit, such as the ratios related to capital structure, asset usage efficiency, financial solvency, profitability, and equity valuation. S2 also lists the definitions of the industry-specific variables. Furthermore, we consider two equity market variables: industry return and industry volatility. We use the standard deviation of the industry compound return 12 months prior to default as the volatility proxy.

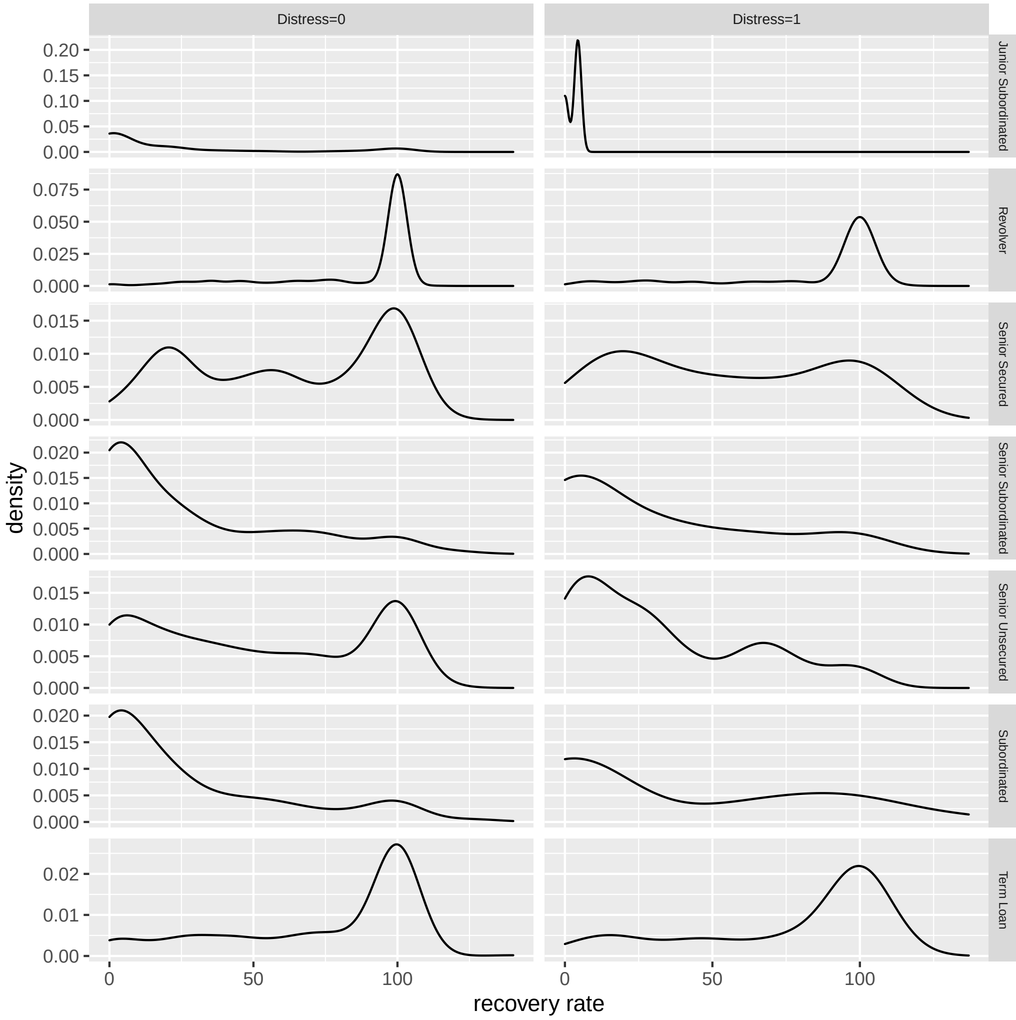

Figure 3. Density estimates of the recovery rate by instrument type and industry distress.

We implement the double-lasso selection with the required variables, $\mathbf{X}_{\text{must}}$ for the instrument type, collateral status, collateral quality ranking, debt cushion, year dummies, industry dummies, and interaction terms of industry distress and instrument types. The to-be-selected variables, $\mathbf{X}_{\text{tobe}}$, are all macroeconomic and industry variables. We have no prior information about the choice of variables in these two categories.

Table 4 lists the variables selected across quantiles. The gray-1 cell indicates that this variable is selected at this quantile, and the nonshaded-0 cell indicates otherwise. We sorted the variable names according to their first appearance in the first quantiles. There are 12 selected variables that are common from the 5th to the 75th quantile. Among the selected variables, seven are macroeconomic and five are industry-specific. Macroeconomic variables include Moody's seasoned Baa corporate bond yield relative to the yield on the 10-year Treasury constant maturity (BAA10Y), total borrowings of depository institutions from the Federal Reserve (BORROW), commercial and industrial loans from all commercial banks (BUSLOANS), corporate profits after tax with inventory valuation adjustment and capital consumption adjustment (CPATAX), change in private inventories (CBI), personal consumption expenditures in durable goods (PCDG), and the S&P 500 return. The industry variables include total debt to EBITDA (debt_ebitda), free cash flow to operating cash flow (fcf_ocf), operating profit margin after depreciation (opmad), industry return, and industry volatility. The price-to-sales ratio (ps) is also common across all quantiles, except the 5th quantile function. The minimum and the maximum numbers of the covariates are 12 and 22 at the 5% and 25% percentiles, respectively. These relatively small numbers show that the majority of the macroeconomic and industry variables are not relevant to identify the effect of the industry distress after conditioning the effects of the required debt characteristics.[6]

Table 4. Variables selected by the lasso-based double selection method.

Percentile	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%
BAA10Y	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BORROW	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BUSLOANS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CBI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CPATAX	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
debt_ebitda	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fcf_ocf	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
industry return	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
industry volatility	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
opmad	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PCDG	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
S&P500	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
cash_ratio	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ps	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
quick_ratio	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
GProf	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
CAPEI	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
staff_sale	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
lt_debt	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0
TEDRATE	0	0	1	0	1	0	0	0	0	0	0	0	1	1	1
CES3000000008	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
DPIC96	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
npm	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
M2SL	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
TOTALSL	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
M1SL	0	0	0	1	1	1	1	0	1	1	1	0	0	0	0
EMRATIO	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
VIXCLS	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
IPFINAL	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
PEG_1yrforward	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1
UEMP5TO14	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
PERMITW	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
MORTG	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
dpr	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
CNCF	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
IPB51200SQ	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
curr_ratio	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
HSN1F	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
USROE	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
debt_at	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
totdebt_invcap	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
HOUSTNE	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
covariates no.	12	16	21	19	22	20	21	17	17	16	18	18	21	21	21

3.3. Unconditional Quantile Regression Estimates

First, we compare the unconditional median regression and the conditional median regression in Table 5. The last column lists the conditional mean estimates from the OLS for comparison. We use the same set of covariates that were selected from the lasso-based double selection approach. The results show that industry distress has a statistically significant and negative effect when we adopt the UQR approach. The UQR estimate is −9.82 with a standard error of 1.00. As the median value of the recovery rate is 65.41%, this result means that when the industry peers are in distress, the median value of the recovery rate lowers to 55.59%. In contrast, the CQR and the OLS produce insignificantly negative estimates. The *p*-value for industry distress in the CQR is 0.67, and the *p*-value for the OLS is 0.22.[7]

Table 5. Unconditional median reg. versus conditional median reg.

	UQR			CQR			OLS		
	Coef.	**s.e.**	***p*-Value**	**Coef.**	**s.e.**	***p*-Value**	**Coef.**	**s.e.**	***p*-Value**
constant	565.02	31.86	0.00	119.12	76.87	0.12	198.88	59.09	0.00
industry distress	−9.82	1.00	0.00	−1.21	2.88	0.67	−3.56	2.91	0.22
percent below	1.10	0.05	0.00	0.39	0.03	0.00	0.41	0.02	0.00
collateral	22.14	1.81	0.00	14.11	4.31	0.00	9.33	2.22	0.00
Junior Subordinated	−30.66	3.85	0.00	−26.19	4.76	0.00	−20.09	4.19	0.00
Senior Secured	−37.24	1.53	0.00	−12.06	2.59	0.00	−12.43	1.69	0.00
Senior Subordinated	−35.85	2.45	0.00	−26.38	4.45	0.00	−18.51	2.65	0.00
Senior Unsecured	−9.07	1.81	0.00	−2.98	4.18	0.48	−2.24	2.37	0.34
Subordinated	−31.00	3.19	0.00	−23.13	4.92	0.00	−16.43	2.80	0.00
Term Loan	−13.44	0.67	0.00	−1.48	1.49	0.32	−6.08	1.38	0.00
Junior Subordinated × distress	−22.31	2.22	0.00	−14.80	10.33	0.15	−15.58	17.03	0.36
Senior Secured × distress	−5.63	3.49	0.11	−1.42	9.79	0.89	−3.52	4.89	0.47
Senior Subordinated × distress	13.58	1.01	0.00	1.39	4.80	0.77	6.35	4.60	0.17
Senior Unsecured × distress	−10.27	1.26	0.00	−5.14	4.10	0.21	−7.04	3.16	0.03
Subordinated × distress	56.38	6.13	0.00	5.93	9.06	0.51	19.45	5.33	0.00
Term Loan × distress	7.37	1.12	0.00	4.36	3.30	0.19	4.61	3.34	0.17
BAA10Y	−3.07	0.94	0.00	−6.10	2.18	0.01	−3.49	1.80	0.05
BORROW	0.14	0.01	0.00	0.06	0.02	0.00	0.06	0.02	0.00
BUSLOANS	0.03	0.01	0.02	0.02	0.02	0.42	0.01	0.01	0.61
CBI	0.17	0.01	0.00	0.00	0.03	0.94	0.03	0.02	0.20
CPATAX	0.22	0.02	0.00	0.08	0.02	0.00	0.08	0.02	0.00
IPB51200SQ	−4.52	0.37	0.00	−0.42	0.81	0.60	−1.04	0.67	0.12
M1SL	0.00	0.01	0.76	0.03	0.02	0.26	0.02	0.02	0.22
PCDG	−0.05	0.01	0.00	−0.01	0.04	0.83	−0.04	0.03	0.16
S&P 500 return	11.75	3.07	0.00	11.38	4.99	0.02	8.28	3.32	0.01
ps	−18.84	1.07	0.00	−8.13	1.60	0.00	−9.54	1.14	0.00
opmad	−158.63	8.94	0.00	−19.14	20.70	0.36	−61.72	14.46	0.00
fcf_ocf	−11.82	1.62	0.00	2.92	5.07	0.56	−0.42	3.67	0.91
debt_ebitda	10.51	0.73	0.00	1.95	1.10	0.08	3.04	0.94	0.00
lt_debt	−230.43	9.81	0.00	−67.74	14.66	0.00	−86.10	10.53	0.00
industry return	6.08	3.45	0.08	−3.75	3.82	0.33	0.23	2.99	0.94
industry volatility	−50.54	3.31	0.00	−4.60	11.08	0.68	−27.94	7.85	0.00
rank dummy	yes			yes			yes		
year dummy	yes			yes			yes		
industry dummy	yes			yes			yes		

Figure 4 highlights the discrepancy between the estimates from the UQR and CQR approaches across quantiles. This figure plots the coefficients and corresponding 95% confidence intervals of industry distress, collateral status, and the percent below for every 5% percentile from the 5th to 75th quantile range. We stop at the 75th quantile, as the UQR estimates do not change after the full recovery (i.e., after the 65th quantile). Table 6 records the coefficients, standard errors, and *p*-values that are consistent with Figure 4. Panel A of Figure 4 displays the "V" shape of the UQR estimates for industry distress. The reduction magnitude of industry distress to recovery rate ratio increases as it moves toward the 40th percentile, but after the 40th percentile, the reduction slows down and even turns positive after the 65th percentile. The industry distress decreases the recovery rates from 2.94% (55th quantile) to 15.80% (40th quantile) in the quantile range from 15th to 55th. The industry distress to recovery rate ratio is significantly positive at the 10th quantile (increase 2.56%) and above the 65th quantile (increase 1.70%). In contrast, the CQR estimates of the industry distress are significantly indifferent to zero, except for the conditional quantiles higher than the 70th one. These results show that the conditional quantiles of the recovery rates are not affected when the issuers' industry peers are in distress or when they are not. We can only interpret the CQR estimates as the weighted average of the predictive effect of distress on the conditional distribution of the recovery rate that is confined to the selected covariates.

Panels B and C of Figure 4 examine the estimates of UQR and CQR for the collateral status and the debt buffer. For the collateral status, both methods show that the recovery rates increase about 8% to 20% if the debt or loan instruments are secured by collateral. The UQR estimates point out that collateral has the most prominent effect on the median of recovery rates (increase 22.14%) and has no significant effect on the 5th quantile. For full recovery, collateral still benefits the recovery by 4.58%. Panel C plots the effect of the debt below on the recovery rate. At the time of default, the percent below is the percentage of the total dollar amount outstanding of all defaulted debt that is contractually subordinated to the current instrument. The percent below provides the buffer for the defaulted instrument and accordingly increases default recovery. The UQR estimate shows that a 1% increase in the debt below makes the recovery rate increase by 1.31% at the 55th quantile and by 0.27% at and above the 65th quantile. The magnitudes of these positive effects differ in different quantiles. These effects increase before the 55th quantile and then reduce afterward. Compared with the CQR estimates, the effect of debt below on the conditional recovery rate is relatively constant.

Table 6. Quantile effects of the industry distress on recovery rate.

	UQR			CQR		
Quantile	**Coef.**	**s.e.**	***p*-Value**	**Coef.**	**s.e.**	***p*-Value**
5%	0.000	0.381	1.000	2.330	4.730	0.622
10%	2.556 ***	0.346	0.000	0.088	5.645	0.988
15%	−3.683 ***	0.326	0.000	−0.076	4.497	0.987
20%	−4.747 ***	0.832	0.000	1.392	3.629	0.701
25%	−8.307 ***	1.714	0.000	−0.968	3.360	0.773
30%	−4.475 **	1.878	0.017	−1.580	3.328	0.635
35%	−13.112 ***	0.970	0.000	0.353	3.351	0.916
40%	−15.798 ***	1.945	0.000	−1.069	3.333	0.748
45%	−13.829 ***	1.632	0.000	−0.923	3.140	0.769
50%	−9.817 ***	1.000	0.000	−1.210	2.877	0.674
55%	−2.943 *	1.716	0.086	−0.638	2.810	0.820
60%	0.342	0.802	0.670	−2.768	2.662	0.298
65%	1.709 ***	0.029	0.000	−1.600	2.614	0.541
70%	1.709 ***	0.029	0.000	−4.532 *	2.513	0.071
75%	1.709 ***	0.029	0.000	−7.657 ***	2.378	0.001

p values less than 0.1, 0.05, and 0.01 are given one, two, and three asterisks respectively.

In Table 7, we present the 12 common covariates selected in the DML-UQR selection process from the 10th to 70th percentiles. In addition to the essential variables we used (instrument type, collateral status, collateral quality ranking, debt cushion, year dummies, industry dummies, and interaction terms of industry distress and instrument types), we also list the estimates of the variables listed in Table 4. Figures 5 and 6 display the UQR estimates of these 12 macroeconomic and industry variables. We found that the percentage below has a small but positive effect on the recovery rate. The collateral dummy has a large and positive effect on the recovery, particularly in the middle quantile, with a coefficient of 22.14 compared with the noncollateral case. Subordination has a negative effect on the recovery rate. The inclusion of selected control variables (such as BAA 10Y, BORROW, BUSLOANS, and other macroeconomic and industry-specific variables) in the DML-UQR is to improve accuracy in identifying and estimating the impact of industry distress (the target variable). As a result, we did not expect all regression quantile coefficients related to these controls to align with the OLS methodology.

Table 8 displays the estimates of quantile-specific covariates in the same quantile range. Our findings indicate that, in addition to the common selected variables, different quantiles require different controls to identify and estimate the impact of industry distress. For instance, in the 10th percentile, we have to include three more industry-specific variables (cash_ratio, GProf, and quick_ratio), and for the 30th percentile, six variables (M1SL, MORTG, UEMP5TO14, VIXCLS, PEG_1yrforward, and sta_sale) are needed. This table illustrates that each quantile has its own unique risk exposures, which require a tailored approach to accurately characterize the effect of industry distress on the recovery rate.

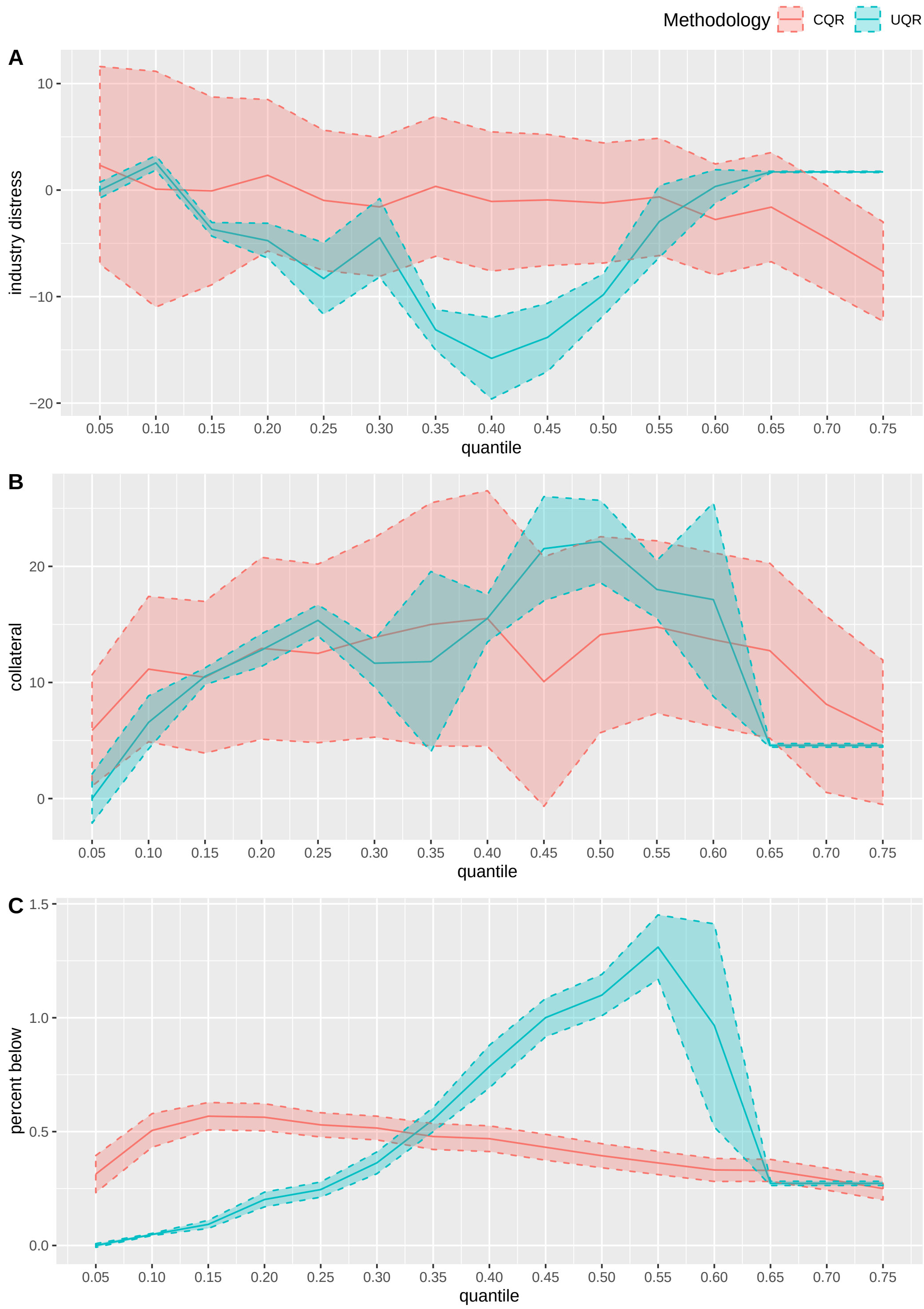

Figure 4. Parameter estimates of industry distress (**A**), collateral (**B**), and percent below (**C**): unconditional quantile regression (UQR) vs. conditional quantile regression (CQR).

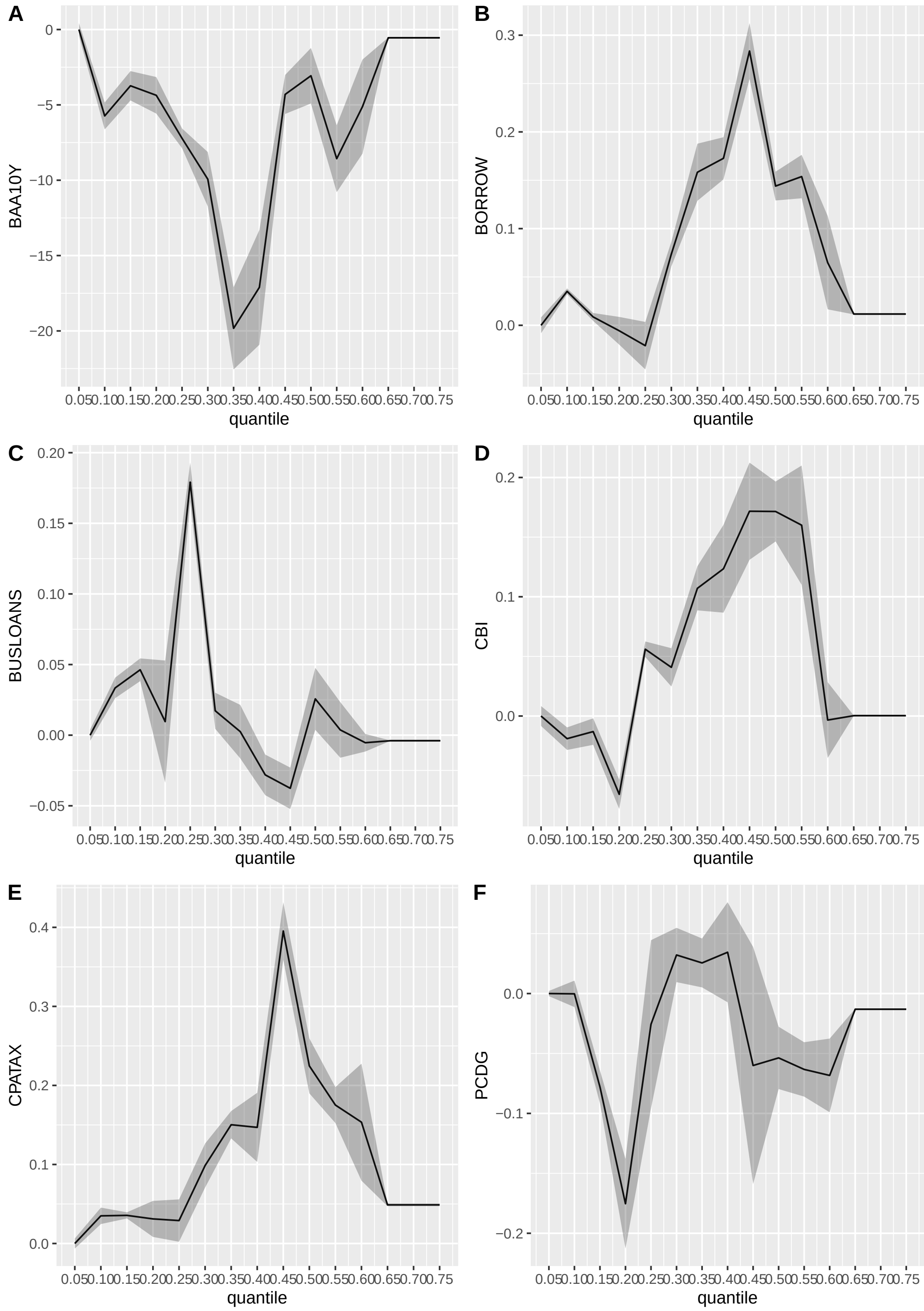

Figure 5. Unconditional quantile regression estimates: macroeconomic variables. (**A**) Moodys seasoned baa corporate bond yield relative to yield on 10-year treasury. (**B**) total borrowings of depository institutions from the Federal Reserve. (**C**) commercial and industrial loans, all commercial banks. (**D**) change in private inventories. (**E**) corporate profits after tax (with IVA and CCAdj). (**F**) personal consumption expenditures: durable goods. Please see online S2 for the description of the variable abbreviations.

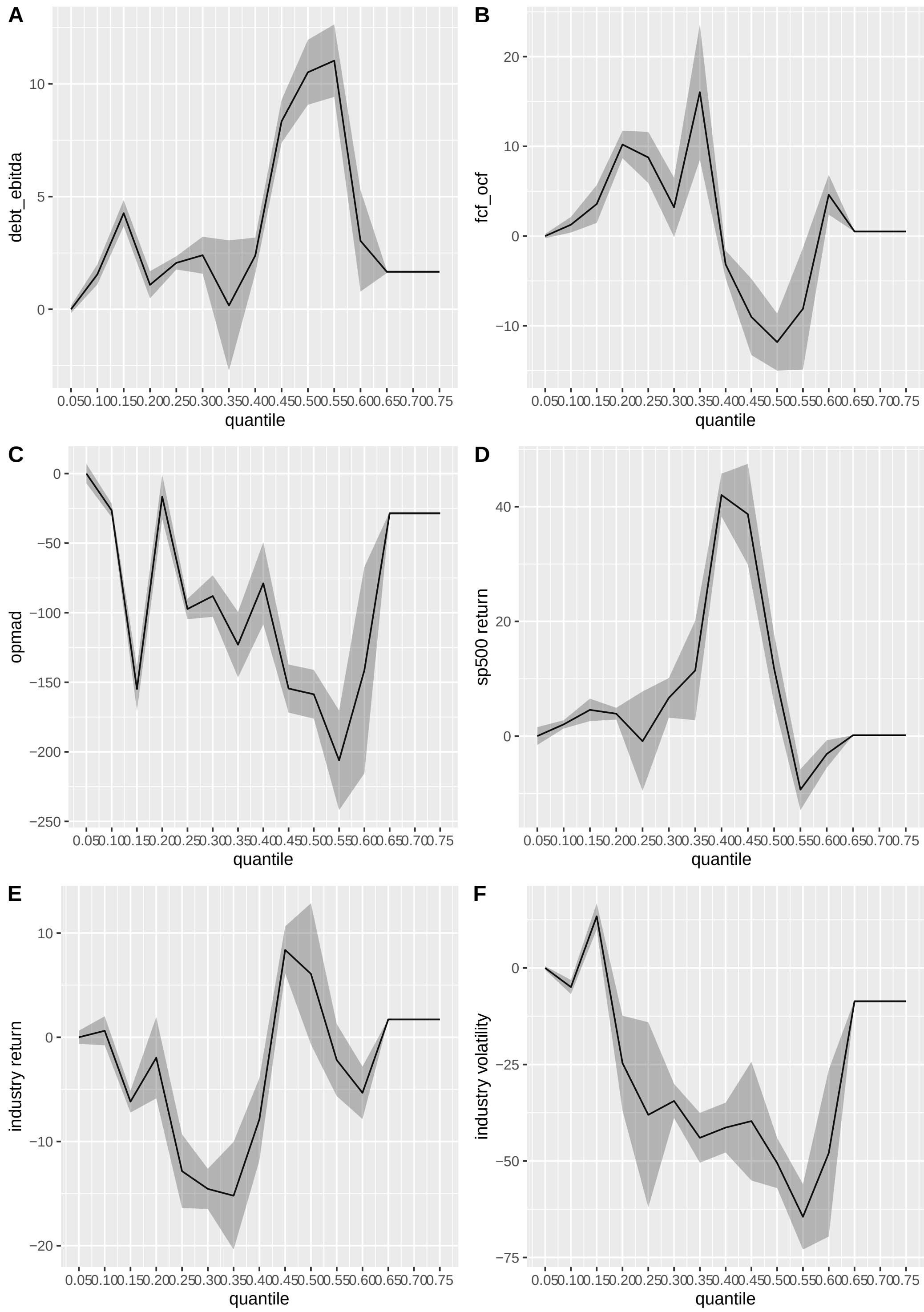

Figure 6. Unconditional quantile regression estimates: industry-specific variables and S&P 500. (**A**) total debt/ebitda. (**B**) free cash flow/operating cash flow. (**C**) operating profit margin after depreciation. (**D**) S&P500 return. (**E**) past 12 month industry return. (**F**) past 12 month standard deviation of the industry return. Please see online S2 for the description of the variable abbreviations.

Table 7. UQR estimates: common covariates.

Percentile	10%		20%		30%		40%		50%		60%		70%	
	Coef.	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**
industry distress	2.56	0.35	−4.75	0.83	−4.48	1.88	−15.80	1.95	−9.82	1.00	0.34	0.80	1.71	0.03
percent below	0.05	0.00	0.20	0.02	0.36	0.02	0.79	0.05	1.10	0.05	0.97	0.23	0.27	0.00
collateral	6.56	1.16	12.79	0.72	11.66	1.04	15.52	1.04	22.14	1.81	17.14	4.26	4.58	0.08
Junior Subordinated	−18.69	1.25	−29.95	2.38	−40.35	1.39	−42.10	2.15	−30.66	3.85	−14.30	2.04	−4.82	0.08
Senior Secured	2.16	0.26	−1.59	1.10	−13.59	0.65	−21.15	1.08	−37.24	1.53	−34.29	6.50	−10.37	0.18
Senior Subordinated	−6.62	0.46	−22.63	2.26	−30.41	1.96	−35.22	1.75	−35.85	2.45	−25.67	5.02	−8.02	0.14
Senior Unsecured	6.86	0.72	5.20	2.28	0.13	0.61	−1.00	0.99	−9.07	1.81	−6.28	0.61	−3.15	0.05
Subordinated	−11.58	0.70	−21.14	3.54	−31.15	1.37	−31.86	1.75	−31.00	3.19	−14.55	2.47	−4.76	0.08
Term Loan	−1.15	0.33	−4.64	0.53	−4.84	0.64	−10.82	0.58	−13.44	0.67	−14.47	3.28	−4.12	0.07
Junior Subordinated × distress	5.42	1.68	−46.22	2.08	−35.79	2.51	−40.19	2.02	−22.31	2.22	−3.99	1.24	−1.75	0.03
Senior Secured × distress	4.21	0.57	3.23	0.89	−3.05	1.46	0.80	1.57	−5.63	3.49	−10.65	3.88	−2.90	0.05
Senior Subordinated × distress	−7.67	0.72	5.64	0.87	11.16	2.14	12.57	2.18	13.58	1.01	9.90	2.52	0.48	0.01
Senior Unsecured × distress	−0.19	0.36	−9.70	1.21	−9.34	1.99	−12.61	0.88	−10.27	1.26	−18.22	2.86	−5.54	0.09
Subordinated × distress	2.97	0.66	5.30	1.45	26.77	2.77	53.36	3.13	56.38	6.13	22.05	9.50	2.87	0.05
Term Loan × distress	1.95	0.45	6.15	0.71	0.48	1.72	13.31	1.53	7.37	1.12	4.52	1.47	1.08	0.02
BAA10Y	−5.73	0.46	−4.36	0.62	−9.93	0.92	−17.11	1.94	−3.07	0.94	−5.12	1.59	−0.55	0.01
BORROW	0.04	0.00	−0.01	0.01	0.07	0.01	0.17	0.01	0.14	0.01	0.06	0.02	0.01	0.00
BUSLOANS	0.03	0.00	0.01	0.02	0.02	0.01	−0.03	0.01	0.03	0.01	−0.01	0.00	0.00	0.00
CBI	−0.02	0.00	−0.07	0.01	0.04	0.01	0.12	0.02	0.17	0.01	0.00	0.02	0.00	0.00
CPATAX	0.03	0.01	0.03	0.01	0.10	0.01	0.15	0.02	0.22	0.02	0.15	0.04	0.05	0.00
PCDG	0.00	0.01	−0.18	0.02	0.03	0.01	0.03	0.02	−0.05	0.01	−0.07	0.02	−0.01	0.00
S&P 500 return	2.03	0.37	3.90	0.53	6.66	1.76	42.03	1.91	11.75	3.07	−3.13	1.21	0.15	0.00
opmad	−26.50	2.51	−16.56	8.12	−88.03	7.62	−78.88	15.05	−158.63	8.94	−141.31	37.84	−28.51	0.49
fcf_ocf	1.27	0.44	10.20	0.78	3.20	1.68	−3.16	0.79	−11.82	1.62	4.61	1.13	0.51	0.01
debt_ebitda	1.55	0.23	1.09	0.31	2.40	0.42	2.38	0.41	10.51	0.73	3.04	1.15	1.66	0.03
ps	−4.76	0.20	−9.32	0.68	−14.08	1.56	−20.87	0.94	−18.84	1.07	−9.25	2.61	−2.56	0.04
industry return	0.62	0.71	−1.96	1.99	−14.55	0.99	−7.88	2.02	6.08	3.45	−5.33	1.29	1.71	0.03
industry volatility	−4.94	0.93	−24.63	6.26	−34.45	2.26	−41.34	3.29	−50.54	3.31	−47.96	11.02	−8.63	0.15
rank	yes		yes		yes		yes		yes		yes		yes	
year	yes		yes		yes		yes		yes		yes		yes	
industry	yes		yes		yes		yes		yes		yes		yes	

Table 8. UQR estimates: quantile-specific covariates.

Percentile	10%		20%		30%		40%		50%		60%		70%	
	Coef.	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**
CES3000000008			5.92	1.33										
DPIC96			0.02	0.00										
HOUSTNE													−0.04	0.00
HSN1F											−0.07	0.01	−0.02	0.00
IPB51200SQ									−4.52	0.37				
M1SL			0.11	0.01	0.08	0.01			0.00	0.01				
M2SL			−0.03	0.01										
MORTG					11.13	1.35								
PERMITW					−0.07	0.01	−0.17	0.01						
TEDRATE													5.43	0.09
TOTALSL			0.13	0.02										
UEMP5TO14					0.02	0.00								
USROE											−4.38	1.01	−1.24	0.02
VIXCLS					−0.61	0.03								
cash_ratio	−0.08	1.67												
curr_ratio											−30.00	7.16	−14.90	0.25
debt_at													−22.42	0.38
dpr							60.47	3.40						
GProf	52.56	3.82												
lt_debt							−147.04	11.76	−230.43	9.81	−146.79	41.71		
PEG_1yrforward					20.09	2.29					21.02	5.20	4.77	0.08
quick_ratio	−3.75	0.79												
staff_sale			223.44	9.14	162.07	13.56	243.80	11.89						
totdebt_invcap													−47.02	0.80

4. Conclusions

Industry distress is an essential factor in determining default recovery rates. In this paper, we use the unconditional quantile regression with lasso-based double selection to explore the effect of industry distress on each quantile of the default recovery rate. We apply the proposed methods to 5334 debt and loan instruments in Moody's Default and Recovery Database from 1990 to 2017. The results show that industry distress induces a 15% to 3% decline in the default recovery rate in different quantiles, except for the extreme tails, in which industry distress has a positive effect. Consequently, the UQR estimates provide policy-relevant and interpretative quantitative measurements to the recovery rates in a downturn required by the Basel Capital Accord.

Supplementary Materials: The following are available at https://www.mdpi.com/article/10.3390/econometrics11010006/s1, S1: Recovery rate by industry, S2: Macroeconomics and industry-specific variables.

Author Contributions: J.-e.C. and H.-C.C. contributed equally to the paper. All authors have read and agreed to the published version of the manuscript.

Funding: Jau-er Chen acknowledges the financial support of the JSPS KAKENHI (Grant No. JP20K01593). Hui-Ching Chuang acknowledges the financial support from the Ministry of Science and Technology (MOST 110-2410-H-155-012) in Taiwan.

Data Availability Statement: Moody's Default and Recovery Database. All data is derived from the Moody's Investors Service proprietary database of issuer, default, and recovery information. Moody's Investors Service analysts use this data to produce the Annual Default Study, read by market participants globally. Clients frequently use this data to conduct credit research as well as build and update credit risk models, particularly those focused on probability of default and loss given default.

Acknowledgments: This paper has benefited from presentation at the annual meeting of the Taiwan Econometric Society. The usual disclaimer applies.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LGD	Loss given default
CQR	Conditional quantile regression
UQR	Unconditional quantile regression
DML	Double machine learning
Lasso	Least absolute shrinkage and selection operator
RIF	Recentered influence function

Notes

1. https://eba.europa.eu/regulation-and-policy/single-rulebook/ (accessed on 1 January 2019).
2. S1 and S2 are at https://github.com/JeC2017/Appendix_Chuang_and_Chen_2023.
3. We implement UQR using the `R` package `uqr`. The default bandwidth is 0.9 times the minimum of the standard deviation and the interquartile range divided by 1.34 times the sample size to the negative one-fifth power; see Silverman (1986).
4. We implement the lasso selection by the `rlasso` function in the `hdm` packages of the `R` program; see Belloni et al. (2014a) for further information.
5. The term spread, 10-year treasury minus 3-month treasury rate used in Krüger and Rösch (2017) is also adopted in Nazemi and Fabozzi (2018).
6. Nazemi and Fabozzi (2018) identify 24 out of 179 macroeconomic variables when applying the lasso approach to the recovery rates of the S&P Capital IQ-similar corporate bond data in the years 2002–2012.
7. We use the bootstrapping method to estimate the standard errors of UQR and CQR coefficients. The bootstrap replication number is set to be 5000. The heteroskedasticity-robust standard errors are used in OLS coefficients. `R` packages `uqr` and `quantreg` estimate the coefficients and standard errors of UQR and CQR; `lmtest` calculates the heteroskedasticity-robust standard errors of OLS estimates.

References

Acharya, Viral V., Sreedhar T. Bharath, and Anand Srinivasan. 2007. Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics* 85: 787–821. [CrossRef]

Altman, Edward, Brooks Brady, Andrea Resti, and Andrea Sironi. 2005. The link between default and recovery rates: Theory, empirical evidence, and implications. *The Journal of Business* 78: 2203–28. [CrossRef]

Bastos, João A. 2010. Forecasting bank loans loss-given-default. *Journal of Banking & Finance* 34: 2510–17.

BCBS. 2004. *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Basel: Basel Committee on Banking Supervision.

BCBS. 2005. *Guidance on Paragraph 468 of the Framework Document*. Basel: Basel Committee on Banking Supervision.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014a. High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives* 28: 29–50. [CrossRef]

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014b. Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies* 81: 608–50. [CrossRef]

Borah, Bijan J. and Anirban Basu. 2013. Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. *Health Economics* 22: 1052–70. [CrossRef]

Bruche, Max and Carlos González-Aguado. 2010. Recovery rates, default probabilities, and the credit cycle. *Journal of Banking & Finance* 34: 754–64.

Chang, Yuanchen, Yi-Ting Hsieh, Wenchien Liu, and Peter Miu. 2020. Intra-industry bankruptcy contagion: Evidence from the pricing of industry recovery rates. *European Financial Management* 26: 503–34. [CrossRef]

Chava, Sudheer, Catalina Stefanescu, and Stuart Turnbull. 2011. Modeling the loss distribution. *Management Science* 57: 1267–87. [CrossRef]

Chen, Jau-er, and Chen-Wei Hsiang. 2019. Causal random forests model using instrumental variable quantile regression. *Econometrics* 7: 49. [CrossRef]

Chen, Jau-er, Chien-Hsun Huang, and Jia-Jyun Tien. 2021. Debiased/Double machine learning for instrumental variable quantile regressions. *Econometrics* 9: 15. [CrossRef]

EBA. 2018. *Final Draft Regulatory Technical Standards on the Specification of the Nature, Severity and Duration of an Economic Downturn in Accordance with Articles 181(3)(a) and 182(4)(a) of Regulation (EU) No 575/2013*. London: European Banking Authority.

EBA. 2019. *Final Report: Guidelines for the Estimation of LGD Appropriate for an Economic Downturn*. London: European Banking Authority.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. Taming the factor zoo: A test of new factors. *The Journal of Finance* 75: 1327–70. [CrossRef]

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica* 77: 953–73.

Gambetti, Paolo, Geneviève Gauthier, and Frédéric Vrins. 2019. Recovery rates: Uncertainty certainly matters. *Journal of Banking & Finance* 106: 371–83.

Gürtler, Marc, and Martin Hibbeln. 2013. Improvements in loss given default forecasts for bank loans. *Journal of Banking & Finance* 37: 2354–66.

Hartmann-Wendels, Thomas, Patrick Miller, and Eugen Töws. 2014. Loss given default for leasing: Parametric and nonparametric estimations. *Journal of Banking & Finance* 40: 364–75.

Jankowitsch, Rainer, Florian Nagler, and Marti G. Subrahmanyam. 2014. The determinants of recovery rates in the US corporate bond market. *Journal of Financial Economics* 114: 155–77. [CrossRef]

James, Christopher, and Atay Kizilaslan. 2014. Asset specificity, industry-driven recovery risk, and loan pricing. *Journal of Financial and Quantitative Analysis* 49: 599–631. [CrossRef]

Kellner, Ralf, Maximilian Nagl, and Daniel Rösch. 2022. Opening the black box—Quantile neural networks for loss given default prediction. *Journal of Banking & Finance* 134: 106334.

Koenker, Roger. 2005. *Quantile Regression*. Cambridge: Cambridge University Press.

Koenker, Roger, and Gilbert Bassett. 1978. Regression quantiles. *Econometrica* 46: 33–50. [CrossRef]

Krüger, Steffen, and Daniel Rösch. 2017. Downturn LGD modeling using quantile regression. *Journal of Banking & Finance* 79: 42–56.

Nazemi, Abdolreza, and Frank J. Fabozzi. 2018. Macroeconomic variable selection for creditor recovery rates. *Journal of Banking & Finance* 89: 14–25.

Nazemi, Abdolreza, Friedrich Baumann, and Frank J. Fabozzi. 2022. Intertemporal defaulted bond recoveries prediction via machine learning. *European Journal of Operational Research* 297: 1162–77. [CrossRef]

Nazemi, Abdolreza, Konstantin Hdidenreich, and Frank J. Fabozzi. 2018. Improving corporate bond recovery rate prediction using multi-factor support vector regressions. *European Journal of Operational Research* 271: 664–75. [CrossRef]

Maclean, Johanna Catherine, Douglas A. Webber, and Joachim Marti. 2014. An application of unconditional quantile regression to cigarette taxes. *Journal of Policy Analysis and Management* 33: 188–210. [CrossRef]

Porter, Stephen R. 2015. Quantile regression: Analyzing changes in distributions instead of means. In *Higher Education: Handbook of Theory and Research*. New York: Springer.

Qi, Min, and Xinlei Zhao. 2011. Comparison of modeling methods for loss given default. *Journal of Banking & Finance* 35: 2842–55.

Sasaki, Yuya, Takuya Ura, and Yichong Zhang. 2022. Unconditional quantile regression with high-dimensional data. *Quantitative Economics* 13: 955–78. [CrossRef]

Shleifer, Andrel, and Robert W. Vishny. 1992. Liquidation values and debt capacity: A market equilibrium approach. *The Journal of Finance* 47: 1343–66. [CrossRef]

Siao, Jhao-Siang, Ruey-Ching Hwang, and Chih-Kang Chu. 2016. Predicting recovery rates using logistic quantile regression with bounded outcomes. *Quantitative Finance* 16: 777–92. [CrossRef]

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. London: Chapman Hall/CRC Monographs.

Somers, Mark, and Joe Whittaker. 2007. Quantile regression for modelling distributions of profit and loss. *European Journal of Operational Research* 183: 1477–87. [CrossRef]