“Threshold Events and Identification: A Study of Cash Shortfalls”

Paper Replication by
Fikrat Valehli

Introduction and Summary of the Usage of Thresholds for Identification

The study by Bakke and Whited 2012, examines the role of threshold events in identifying corporate strategies. Their identification approach is similar to the Regression Discontinuity Design (RDD). The threshold event used in the study is the mandatory contributions that firms must make to their defined benefit (DB) pension plans. “These contributions are a direct function of the level of pension plan funding (assets minus liabilities), and they jump up discretely at several different funding thresholds.”

The authors also examine whether the cash shortfalls after the DB events lead to a decline in investment. The endogeneity concern is that the mandatory pension contributions may be endogenous and related to capital expenditures and investment activities. However, there is an element of randomness in the firm’s obligation to pay for pension contributions. The gap between the pension assets and liabilities, which is the market value of these contributions, creates threshold events that make it mandatory for firms to contribute to the pension plan. Because the government rule that determines these threshold values consists of kinks and the market value of pension assets is subject to exogenous fluctuations, the firm’s ability to manipulate the funding gap through pension contributions or changes in employment and plan benefits is limited. Therefore, the authors argue that near the threshold, some firms are exogenously assigned to the treatment group, they are obligated to pay benefits, and some are not. Therefore, by comparing their capital expenditures, we might be able to identify the impact of cash shortfalls on investment activities.

Rauh (2006) initially attempts to identify the causal relationship between mandatory contributions in the DB pension plans and investment decisions using the RDD approach. However, Bakke and Whited (2012) are puzzled by the magnitude of the effects found in Rauh (2006) because pension contributions are relatively small compared to investment decisions. This paper attempts to understand the magnitude of the impact.

Methodology

My replication of Bakke and Whited (2012) consists of six steps.

1. I find pension plans data from the Center for Retirement Research of Boston College.[1] The IRS 5500 forms are posted on the official IRS website starting from 1999. However, to be able to match with CUSIPs of the firms, the authors conduct the study with a sample spanning 1990 to 1998, and the data can be found on the above-mentioned website.
   The IRS-5500 forms datasets are gigantic, and it requires deletion of many unnecessary columns for the study. The columns utilized in the study are as follows.

   'TYPE_PENSION_BENEFIT_IND' – shows if the pension plans are DB.
   'SPONS_DFE_EIN' – EIN of the firms.
   'SPONS_DFE_PN' – number of the specific pension plan belong to the firm.
   'CUSIP_ISSUER_NUM' – CUSIP number of the firm, used for matching.
   'TOT_ASSETS_BOY_AMT' – pension plan assets beginning of year (BoY)
   'TOT_LIABILITIES_BOY_AMT' – pension liabilities BoY
   'TOT_DISTRIB_BNFT_AMT' – normal cost of the plan
   'TOT_CONTRIB_AMT' – total contributions to the plan

   Each firm may sponsor more than one pension plan. SPONS_DFE_PN shows the unique number of each pension plan give the firm EIN – SPONS_DFE_EIN.

2. I downloaded the Compustat files. The item list to download is attached in this working directory under compustat_variables.xslx name. The Compustat files contain the firm specific variables, describing their market-to-book ratio, nonpension contributions, capital expenditures, total assets, etc.

3. Match Compustat and IRS 5500 files year by year by CUSIP number. Authors indicate that they also used EIN and firm names while matching. But since they do not indicate how I only use CUSIP for matching. Almost all of the firms in IRS and Compustat data have EIN hence matching by EIN would increase the sample size multiple times. However, without EIN matching, my sample size is only 70% of the authors' sample size. Therefore, I suppose authors matched with EIN under some circumstances. Since I do not know the circumstances, I used only CUSIPs.

4. In order to determine the mandatory contributions, I created deficit reduction contribution (DRC) and minimum funding contribution (MFC) variables following the author's methodology on page 1093. Furthermore, I follow Rauh (2006) when creating Funding Status (FS) variable. I average the funding gaps across the pension plans, thus getting a single average funding quantity per firm.

[1] https://crr.bc.edu/data/form-5500-annual-reports/

5. I merge cross-sectional data over the years to create the panel. The panel is unbalanced, and the sample size is (n * t) = 5424. I suppose it is less than the author's sample because I did not match EIN and Corporate Names.

6. I report the panel data estimates with fixed time and firm effects and export descriptive statistics for the panel. I follow Rauh's (2006) panel data estimation. Moreover, to understand the drivers of the results, I follow Bakke and Whited (2012) and report descriptive statistics for varying subsamples. These subsamples contain the firms that are underfunded in their pension plan, meaning that pension plan assets exceed liabilities. I further export the descriptive statistics for firms that are 90 and 80 percent underfunded.

A clear shortcoming in my approach is not matching firms also by EIN. The main problem is, in the sample, some firms did not provide CUSIP numbers for some pension plans that they sponsor. As a result of my matching by only EIN, my sample sometimes contains only some pension plans of the firms. Moreover, running the regression in some subsamples would help my method to clearly see where the identification coming from.

Results and Comparison

Panel OLS Estimation Summary

Dep. Variable:	inv_to_assts	R-squared:	0.0016
Estimator:	PanelOLS	R-squared (Between):	0.0183
No. Observations:	3877	R-squared (Within):	0.0017
Date:	Fri, Apr 28 2023	R-squared (Overall):	0.0190
Time:	14:43:45	Log-likelihood	8659.2
Cov. Estimator:	Unadjusted		
		F-statistic:	1.2101
Entities:	660	P-value	0.3043
Avg Obs:	5.8742	Distribution:	F(4,3106)
Min Obs:	1.0000		
Max Obs:	9.0000	F-statistic (robust):	1.2101
		P-value	0.3043
Time periods:	108	Distribution:	F(4,3106)
Avg Obs:	35.898		
Min Obs:	2.0000		
Max Obs:	311.00		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI

MTB_ratio	0.0001	0.0005	0.2767	0.7821	-0.0009	0.0012
nonpension_cash	0.0131	0.0066	1.9964	0.0460	0.0002	0.0260
fs_to_assets	-5.974e-09	8.23e-09	-0.7259	0.4680	-2.211e-08	1.016e-08
mc_to_assts	1.388e-11	4.969e-10	0.0279	0.9777	-9.604e-10	9.882e-10

==========

F-test for Poolability: 9.3906 | P-value: 0.0000 | Distribution: F(766,3106)
Included effects: Entity, Time

Table 1: Replication estimates of Bakke and Whited (2012)

Market-to-Book	0.019	0.019	0.019	0.019	0.019	0.019
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Nonpension	0.113	0.113	0.112	0.111	0.111	0.112
Cash Flow	(0.011)	(0.011)	(0.008)	(0.008)	(0.009)	(0.011)
MPCs	−0.638	−0.624				
	(0.267)	(0.268)				
Funding	0.024					
Status	(0.011)					
Funding		0.038	0.048			
Gap		(0.020)	(0.020)			
Violation			−0.020			
Indicator (×10)			(0.014)			
Distance from				0.053		
90% Underfunding				(0.021)		
90% Underfunding				−0.029		
Indicator (×10)				(0.014)		
Distance from					0.058	
80% Underfunding					(0.021)	
80% Underfunding					−0.022	
Indicator (×10)					(0.016)	
Distance from						0.030
the Kink						(0.030)
Kink Indicator						0.020
(×10)						(0.014)
R^2 (within)	0.101	0.101	0.100	0.100	0.100	0.100
R^2	0.684	0.684	0.684	0.684	0.684	0.691

Table 2: Bakke and Whited (2012). Panel data estimations

My replication only corresponds to the first column of the Bakke and Whited (2012) results. The variable fc_to_assets is the Funding Status variable and mc_to_assets is the MPC variable in the paper. Even though the direction of my results is the same as Bakke and Whited's results, unlike the authors, only the non-pension cashflow is the significant variable in my replication. The mandatory pension contributions have a negative and significant impact on the investment variable in the authors' paper. My replication finding shows that it is negative but not significant. Due to the time limitations, I did not run the other regressions with different definitions of the funding gap. The definition I used for the Funding Status variable was the average funding status described above.

Discussion and Endnotes

Bakke and Whited (2012) raise their doubts about the identification strategy of Rauh's (2006) paper, indicating that the RDD setup used in the paper does not necessarily identify the causality between pension contributions and investment. Furthermore, they indicate that considering the relatively small magnitude of the cash shortfall created by the mandatory pension contributions,

the large magnitude of Rauh's (2006) results is puzzling. In fact, the authors attempt to argue that the identification of Rauh (2006) will be valid only if the cash shortfalls resulting from the mandatory contributions has a significant amount. Authors claim that they "find that the strong sensitivity of investment to mandatory contributions stems from heavily underfunded firms that constitute a small fraction of the sample and that are different from the rest of the sample in important ways, that is, the control group differs from the treated group" (Bakke and Whited, 2012).

However, Bakke and Whited (2012) are not very successful in showing that this criticism is valid. The authors run kernel regression of the investment decisions on the mandatory contributions and find that the most significant results are for the 8% and 27% underfunded firms. But these estimates become insignificant when they use clustered standard errors – which is necessary if one wants to get asymptotically unbiased variance estimation. Nonetheless, the direction of Bakke and Whited's (2012) criticism is valid, because if the cash shortfall resulting from the MPC is small, the dosage of treatment is small, thus the results are not expected to be pronounced.

Moreover, Bakke and Whited (2012) do not mention one criticism, which could relevant. Assigning firms to the treatment and control group based on the MPC does not mean that the potential outcomes of the firms are similar. In other words, being close to the MPC cutoff value does not mean that the firm is identical to the firm on the other side of the threshold. However, this is at the core of the RDD identification assumption.

References

BAKKE, T.-E. and WHITED, T.M. (2012), Threshold Events and Identification: A Study of Cash Shortfalls. The Journal of Finance, 67: 1083-1111. https://doi.org/10.1111/j.1540-6261.2012.01742.x

Rauh, Joshua, 2006, Investment and financing constraints: Evidence from the funding of corporatepension plans,Journal of Finance61, 33–71.

Appendix

Compared statistics for the different levels of funding.

less than 80% coverage	book assets	capital expenditures	the book value of common equity	cash	the number of common shares
count	5240.0	4548.0	5238.0	5154.0	5273.0
mean	5270.6	221.2	983.3	375.0	61.6
std	19437.9	702.9	2517.4	2238.0	159.2
less than 90% coverage	book assets	capital expenditures	the book value of common equity	cash	the number of common shares
count	5240.0	4548.0	5238.0	5154.0	5273.0

mean	5270.6	221.2	983.3	375.0	61.6
std	19437.9	702.9	2517.4	2238.0	159.2
less than 1000% coverage	**book assets**	**capital expenditures**	**the book value of common equity**	**cash**	**the number of common shares**
count	5241.0	4549.0	5239.0	5155.0	5274.0
mean	5269.6	221.2	983.1	375.0	61.6
std	19436.2	702.8	2517.1	2237.8	159.2
Whole sample	**book assets**	**capital expenditures**	**the book value of common equity**	**cash**	**the number of common shares**
count	5272.0	4580.0	5270.0	5186.0	5305.0
mean	5272.2	221.6	988.0	373.6	61.8
std	19386.7	701.5	2516.4	2231.3	159.2
less than 80% coverage	**long-term debt**	**employment**	**inventories**	**equity issuance**	**the share price**
count	5200.0	4817.0	5047.0	4467.0	4972.0
mean	959.0	14.4	407.5	27.4	28.6
std	4374.4	37.3	2424.3	136.0	21.5
less than 90% coverage	**long-term debt**	**employment**	**inventories**	**equity issuance**	**the share price**
count	5200.0	4817.0	5047.0	4467.0	4972.0
mean	959.0	14.4	407.5	27.4	28.6
std	4374.4	37.3	2424.3	136.0	21.5
less than 1000% coverage	**long-term debt**	**employment**	**inventories**	**equity issuance**	**the share price**
count	5201.0	4818.0	5048.0	4468.0	4973.0
mean	958.8	14.4	407.4	27.4	28.6
std	4374.0	37.3	2424.0	136.0	21.5
Whole sample	**long-term debt**	**employment**	**inventories**	**equity issuance**	**the share price**
count	5232.0	4849.0	5079.0	4498.0	4994.0
mean	961.2	14.5	408.3	27.3	28.6
std	4364.6	37.4	2417.4	135.5	21.6
less than 80% coverage	**capital expenditure to asset ratio**	**nonpension cash to assets**	**mc_to_assts**	**fs_to_assets**	**market to book value**
count	4547.0	4564.0	5150.0	5237.0	4798.0
mean	0.1	0.1	-97225.6	74133.3	0.2
std	0.0	0.1	#########	117404.2	1.0
less than 90% coverage	**capital expenditure to asset ratio**	**nonpension_cash_to_assts**	**mc_to_assts**	**fs_to_assets**	**market to book value**
count	4547.0	4564.0	5150.0	5237.0	4798.0
mean	0.1	0.1	-97225.6	74133.3	0.2
std	0.0	0.1	#########	117404.2	1.0

less than 1000% coverage	capital expenditure to asset ratio	nonpension_cash_to_assts	mc_to_assts	fs_to_assets	market to book value
count	4548.0	4565.0	5151.0	5238.0	4799.0
mean	0.1	0.1	-97206.2	74120.3	0.2
std	0.0	0.1	#########	117396.8	1.0
Whole sample	capital expenditure to asset ratio	nonpension_cash_to_assts	mc_to_assts	fs_to_assets	market to book value
count	4579.0	4595.0	5159.0	5267.0	4820.0
mean	0.1	0.1	-97035.3	73842.0	0.2
std	0.0	0.1	#########	117167.4	1.0

	Full Sample	In Violation	Not in Violation	<90% Funded	<80% Funded
Total assets	3,418	3,435	3,409	2,362	2,152
Average Funding Status Indicator	0.218	0.622	0.000	0.951	0.970
Plan Violation Indicator	0.351	1.000	0.000	1.000	1.000
Average Funding Status	0.036	0.005	0.052	−0.020	−0.025
Funding Gap	0.015	−0.012	0.030	−0.025	−0.031
Total Contributions	0.003	0.005	0.002	0.008	0.008
Pension Assets	0.143	0.143	0.143	0.115	0.090
Pension Liabilities	0.106	0.138	0.089	0.137	0.125
Mandatory Contributions	0.001	0.003	0.000	0.005	0.006
Investment	0.069	0.066	0.071	0.057	0.055
Cash Flow	0.096	0.089	0.100	0.068	0.057
Nonpension Cash Flow	0.100	0.095	0.102	0.077	0.067
Market-to-Book	1.481	1.460	1.493	1.363	1.346
R&D	0.017	0.018	0.016	0.016	0.016
Advertising	0.012	0.013	0.012	0.010	0.009
Debt-to-Assets	0.249	0.263	0.241	0.280	0.284
Bond Rating	0.411	0.400	0.417	0.271	0.193
Short Term Debt Issuance	0.005	0.004	0.006	0.005	−0.001
Long Term Debt Issuance	0.021	0.022	0.021	0.021	0.027
Saving	0.004	0.004	0.004	0.003	0.003
Cash	0.070	0.065	0.073	0.072	0.076
Dividends	0.019	0.015	0.021	0.010	0.011
Common Dividends per Share	0.615	0.482	0.688	0.296	0.292
Equity Issuance	0.012	0.013	0.012	0.015	0.016
Equity Repurchases	0.012	0.012	0.012	0.007	0.007
Employment % Change	0.808	0.520	0.965	−1.691	−2.853
Earnings	0.042	0.034	0.046	0.014	0.006
Z-Score	2.780	2.280	3.052	1.667	1.599