Exporing Discretionary Accruals: A Stub Project*

Joachim Gassen
TRR 266 Accounting for Transparency
Humboldt-Universität zu Berlin
gassen@wiwi.hu-berlin.de

Ulrich Atz
New York University
Stern School of Business
ua6@stern.nyu.edu

May 11, 2021

Abstract

The Open Science movement promotes the accessibility and reusability of research. This repository has the objective to help researchers establishing such an collaboration-oriented workflow. It uses a toy project on discretionary accruals for demonstration.

*I acknowledge financial support from Deutsche Forschungsgemeinschaft-Project-ID 403041268 - TRR 266.

1 Introduction

This is not a paper but a stub that is part of a project template repository. We developed this repository to provide a prototype for a reproducible and collaborative workflow. Several authors have discussed advantages of such workflows (Wilson et al. (2017), Gertler, Galiani, and Romero (2018), Christensen, Freese, and Miguel (2019)) and many scholars across fields have voiced the need for increased reproducibility in science (e.g., Ioannidis (2005), Gelman and Loken (2014), Duvendack, Palmer-Jones, and Reed (2017)).

2 Discretionary Accruals

To demonstrate our workflow, we explore discretionary accruals across the U.S. We calculate modified Jones and Dechow and Dichev type accruals and show their distributional properties. The main purpose of all this, however, is to provide a toy use case for our project template directory that contains all the code to obtain the data, run the analysis and prepare a paper as well as a presentation.

Table 1 presents our data that is based on a simple WRDS pull of Compustat data with financial firms (SIC 6XXX) excluded. We require data to calculate all variables and this drastically reduces the sample size. Modified Jones discretionary accruals are calculated loosely based on Hribar and Nichols (2007) and Dechow and Dichev discretionary accruals are calculated based on (big surprise) Dechow and Dichev (2002). As you will see from 1, discretionary accruals are very noisy constructs, even after limiting the sample to observations with complete data and winsorizing all data to the top and bottom percentile for each year. Figure 2 shows a very prominent heteroscedasticity of discretionary accruals with regards to size. While researchers have tried to address this problem, the distributional properties of these constructs significantly complicate the interpretation of discretionary accrual-related findings. Especially in high powered settings, the measurement error, being highly correlated with size, will tend to load on variables that are unrelated to the underlying economic construct but correlated with size. Table 2 shows some correlations and 3 shows some completely pointless regressions.

3 Discretionary Accruals, revisited

The shortcomings of the traditional accrual models are well documented, for example, Ball (2013) questions whether the standard models for discretionary accruals measure what they promise to do. Recently, scholars have proposed a Bayesian methodology (Breuer and Schütt 2019) to address some aspects (such as extreme estimates and false positives) of this criticism. The core idea, "borrowing strength" as statisticians call it, i.e. shrinking extreme parameters toward a more reasonable grand mean, may also be achieved be classic frequentist methods such as hierarchical models (also known as multilevel or mixed model). They come at a computational cost that are orders of magnitude cheaper.

Figure 3 shows how the two model approaches compare in their size of coefficients for each industry-year. As discussed, the larger outliers in the traditional approach (one regression for each group) are subdued in the hierarchical approach. This desirable feature comes with a limitation. We are only estimating a dozen parameters instead of more than a thousand, so that the in-sample predictions are perhaps worse. Figure 4 and 5 show how the modified Jones model agrees with the hierarchical Jones model in their point estimates for discretionary accruals. Contrary to intuition perhaps, the hierarchical model exhibits larger extreme values, likely because of the smaller number of parameters. We speculate this could also be due to a convergence issue of the model. The R package we employed is `lme4` (Bates et al. 2015).

Firm-level estimates were often not available because they require a sufficient number of observations per firm. With the hierarchical model we mitigate this issue. Figure 6 shows how the firm-level model estimates compars in their size of coefficients to the classic modified Jones model. Again, the larger outliers in the traditional approach (one regression for each group) are subdued in the hierarchical approach. The point estimates are also larger than before (7 and 8). Notice, however, how the interquartile range shrunk for each year in our sample.

Finally, it may be possible to combine the benefits of a hierarchical approach with the benefits of estimating many parameters in a Bayesian model that specifies weak prior information. We outline this approach in the code, but do not present the results as they require further fine-tuning to be interpretable.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

4 Conclusion

Isn't that wonderful? Discretionary accruals rock but what rocks even more is open science and a collaborative workflow. Clone or fork this repository to kickstart your own projects. If you do not like R, consider contributing code in your favorite statistical programming language to the repo. Thanks for reading and enjoy!

References

Ball, Ray. 2013. "Accounting Informs Investors and Earnings Management Is Rife: Two Questionable Beliefs." *Accounting Horizons* 27 (4): 847–53. https://doi.org/10.2308/acch-10366.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. https://doi.org/10.18637/jss.v067.i01.

Breuer, Matthias, and Harm H. Schütt. 2019. "Accounting for Uncertainty: An Application of Bayesian Methods to Accruals Models." Rochester, NY. https://doi.org/10.2139/ssrn.3417406.

Christensen, Garret, Jeremy Freese, and Edward Miguel. 2019. *Transparent and Reproducible Social Science Research: How to Do Open Science.* University of California Press.

Dechow, Patricia M., and Ilia D. Dichev. 2002. "The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors." *The Accounting Review* 77 (s-1): 35–59. https://doi.org/10.2308/accr.2002.77.s-1.35.

Duvendack, Maren, Richard Palmer-Jones, and W. Robert Reed. 2017. "What Is Meant by Replication and Why Does It Encounter Resistance in Economics?" *American Economic Review* 107 (5): 46–51. https://doi.org/10.1257/aer.p20171031.

Gelman, Andre, and Eric Loken. 2014. "The Statistical Crisis in Science." *American Scientist* 102 (6): 460–65. https://doi.org/10.1511/2014.111.460.

Gertler, Paul, Sebastian Galiani, and Mauricio Romero. 2018. "How to Make Replication the Norm." *Nature* 554: 417–19. https://doi.org/10.1038/d41586-018-02108-9.

Hribar, Paul, and Craig D. Nichols. 2007. "The Use of Unsigned Earnings Quality Measures in Tests of Earnings Management." *Journal of Accounting Research* 45 (5): 1017–53. https://doi.org/https://doi.org/10.1111/j.1475-679X.2007.00259.x.

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLOS Medicine* 2 (8): 696–701. https://doi.org/10.1371/journal.pmed.0020124.

Wilson, Greg, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Nederbragt, and Tracy K. Teal. 2017. "Good Enough Practices in Scientific Computing." *PLOS Computational Biology* 13 (6): 1–20. https://doi.org/10.1371/journal.pcbi.1005510.

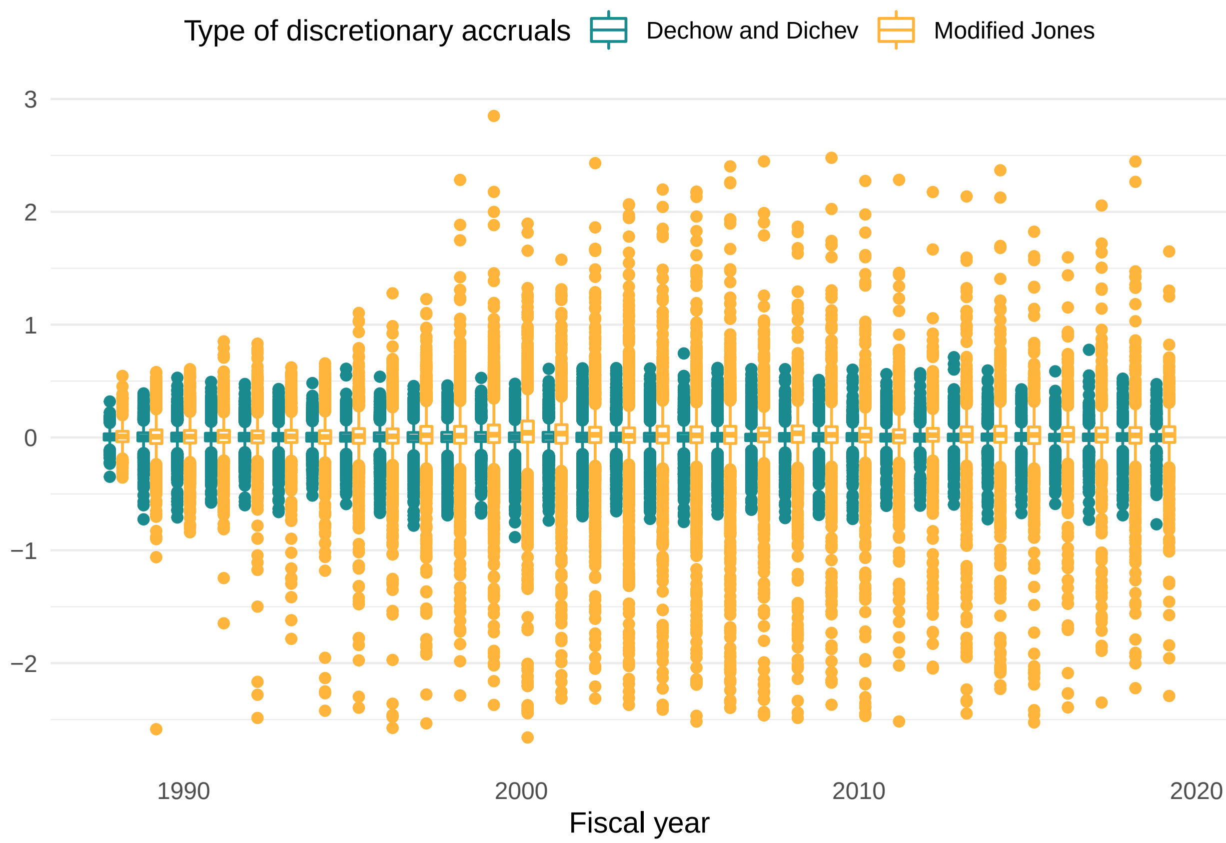

Figure 1: Distribution of Discretionary Accruals over Time

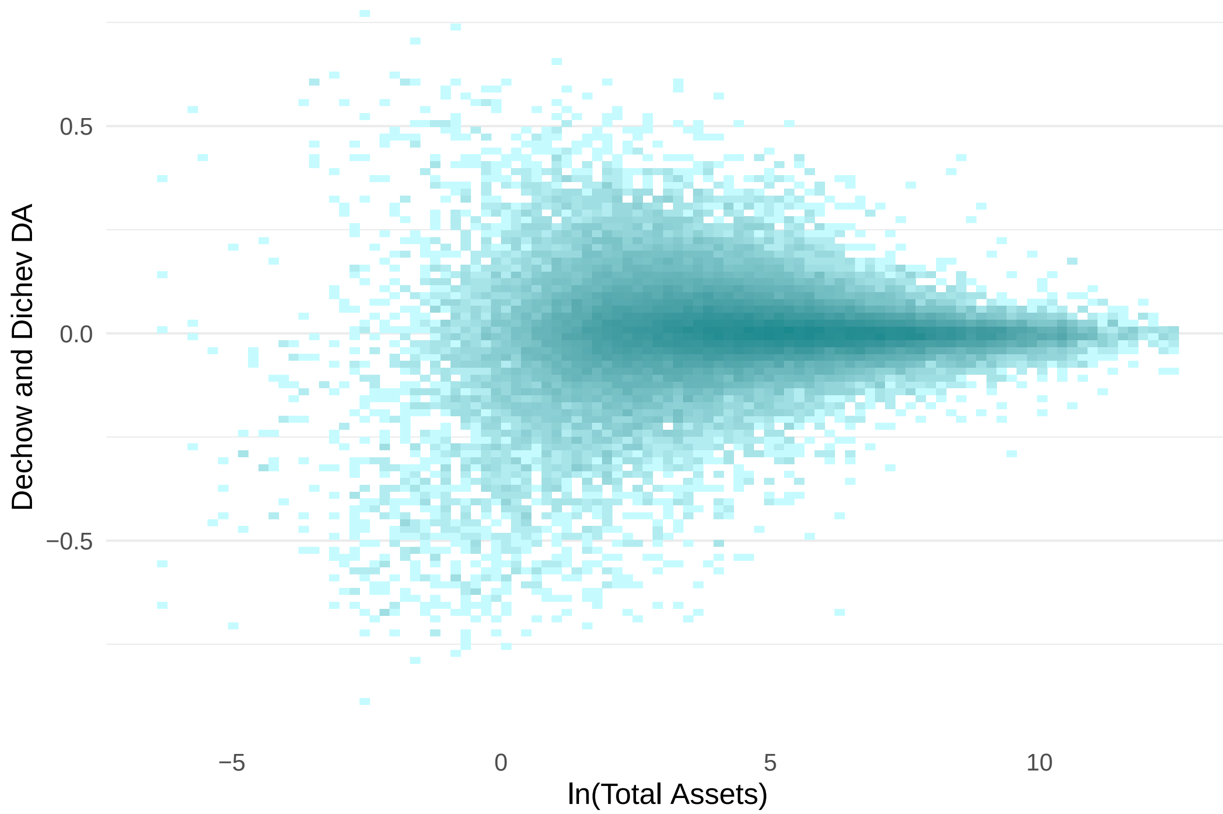

Figure 2: Dechow and Dichev DA and Firm Size

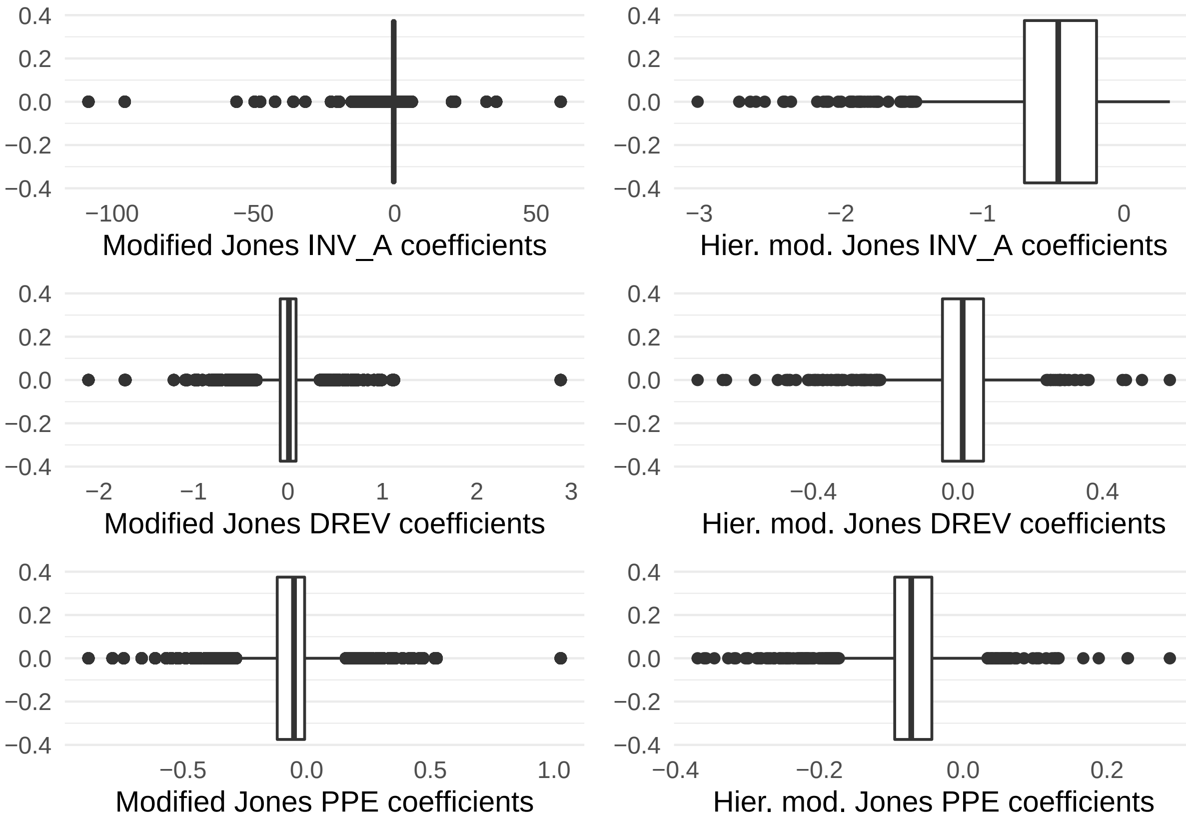

Figure 3: Regression coefficients for the modified Jones model and its hierarchical equivalent (industry-year)

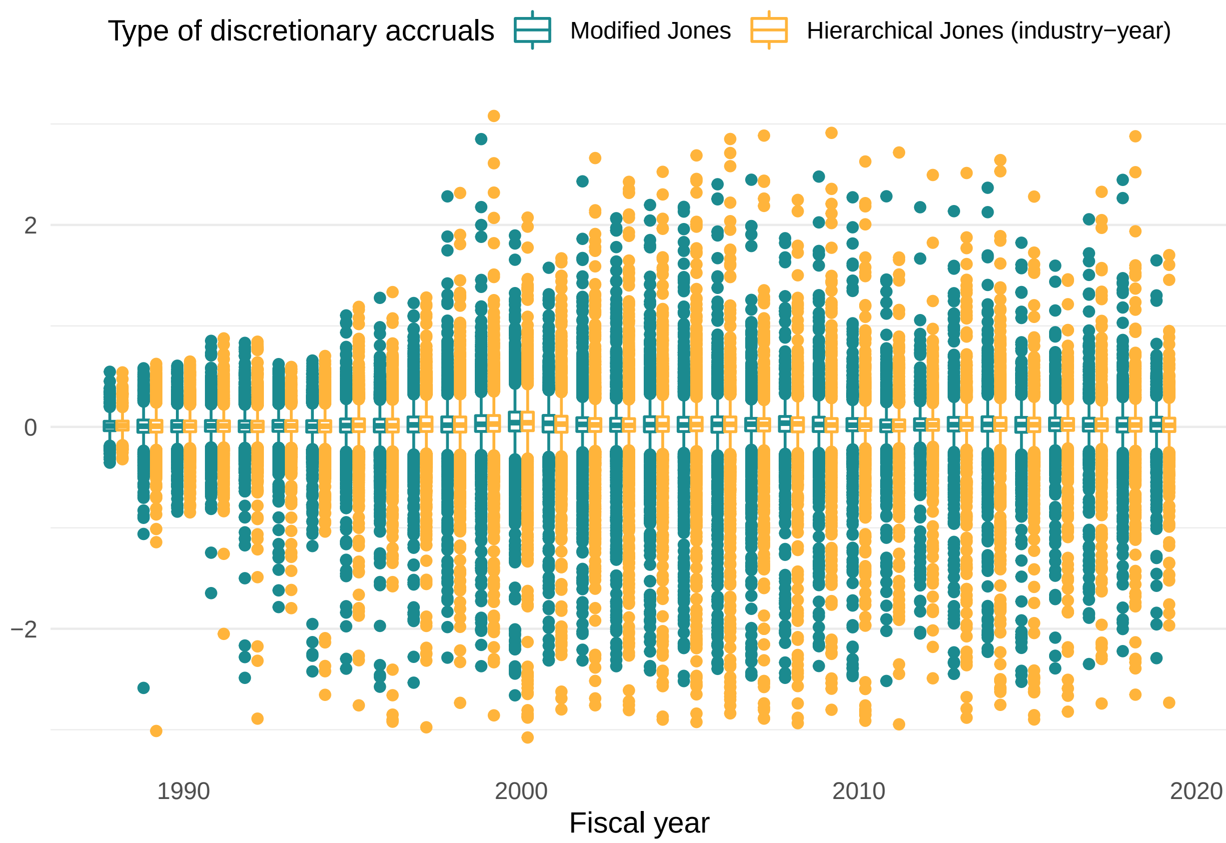

Figure 4: Distribution of Discretionary Accruals over Time, revisited

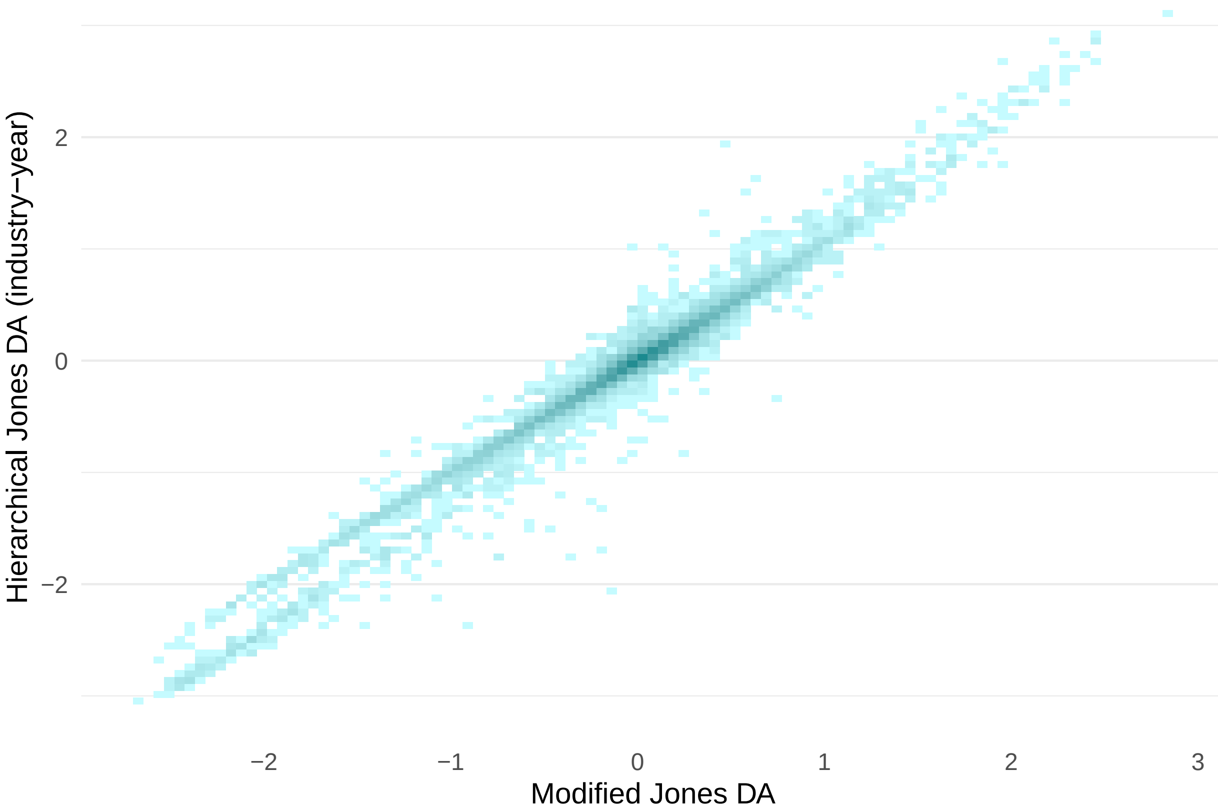

Figure 5: Correlation between the modified Jones model and its hierarchical equivalent

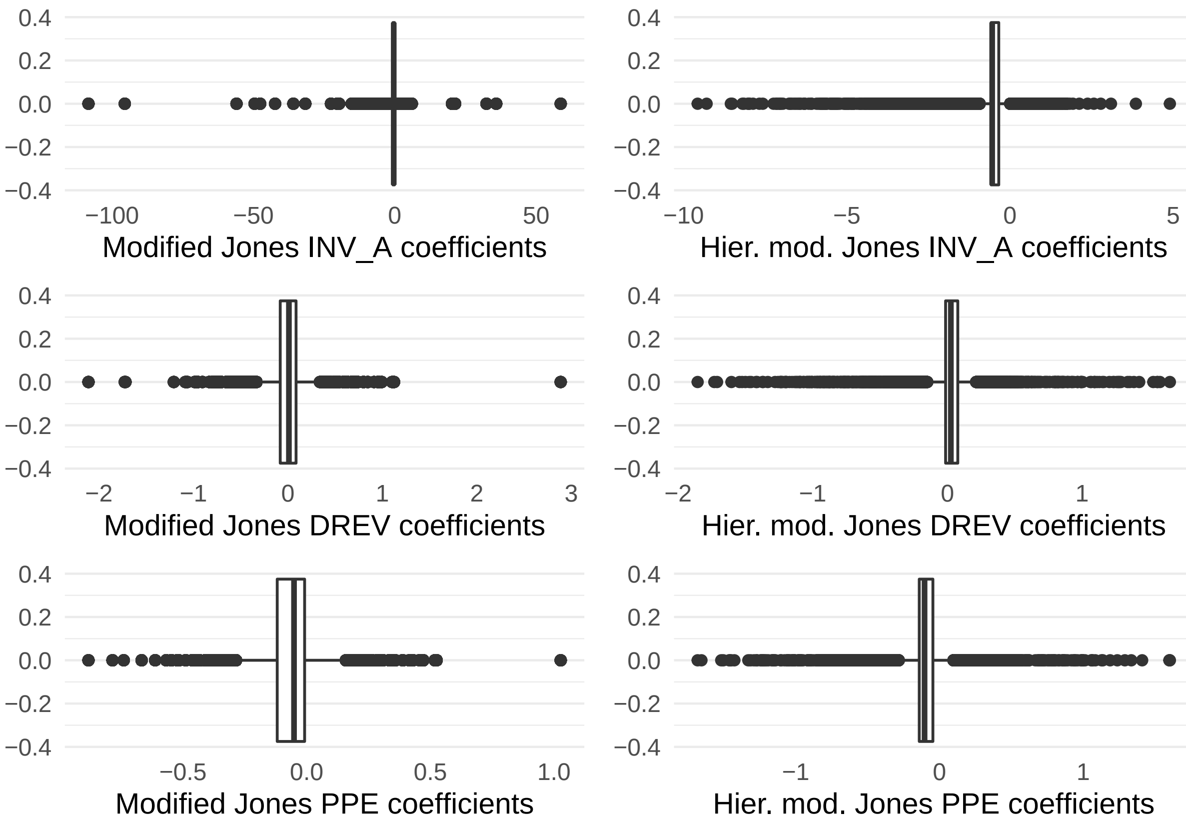

Figure 6: Regression coefficients for the modified Jones model and its hierarchical equivalent (firm-level)

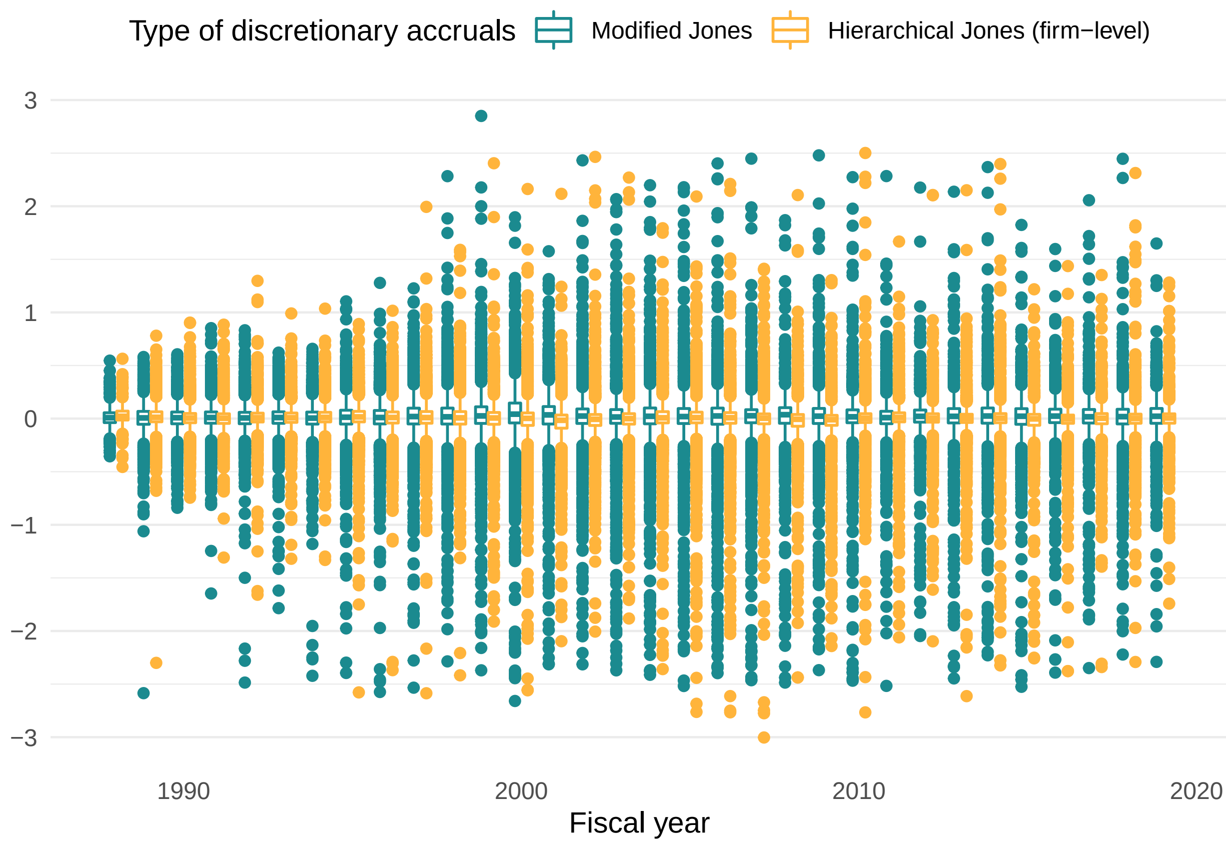

Figure 7: Distribution of Discretionary Accruals over Time, revisited (firm-level)

Figure 8: Correlation between the modified Jones model and its hierarchical equivalent (firm-level

Table 1: Descriptive Statistics

	N	Mean	Std. dev.	Min.	25 %	Median	75 %	Max.
Modified Jones DA	56,224	0.005	0.274	-2.659	-0.047	0.017	0.087	2.850
Dechow and Dichev DA	56,224	0.001	0.104	-0.883	-0.032	0.002	0.038	0.777
Ln(Total assets)	56,224	4.549	2.428	-6.215	2.864	4.453	6.191	12.566
Ln(Market capitalization)	56,224	4.534	2.454	-11.513	2.782	4.421	6.258	14.414
Market to book	56,224	4.991	488.022	-14,752.235	0.929	1.797	3.463	113,537.700
Return on assets	56,224	-0.172	2.510	-362.273	-0.101	0.042	0.088	151.000
Sales growth	56,224	0.968	6.188	-738.707	0.988	1.014	1.056	325.975

Note: The data is obtained from the Compustat U.S. as provided by WRDS. The sample covers the period 1988 to 2019 and 8,781 unique firms.

Table 2: Correlations

	A	B	C	D	E	F	G
A: Modified Jones DA		**0.35**	**0.02**	0.00	-0.01	**0.19**	**-0.01**
B: Dechow and Dichev DA	**0.43**		**0.07**	**0.10**	0.01	**0.11**	**-0.01**
C: Ln(Total assets)	**-0.05**	**-0.01**		**0.88**	-0.00	**0.15**	**0.01**
D: Ln(Market capitalization)	**-0.02**	**0.05**	**0.88**		0.00	**0.08**	**0.01**
E: Market to book	**0.07**	**0.13**	**0.13**	**0.41**		-0.00	-0.00
F: Return on assets	**0.28**	**0.19**	**0.38**	**0.36**	**0.19**		0.00
G: Sales growth	**0.10**	**0.22**	0.00	**0.10**	**0.21**	**0.21**	

This table reports Pearson correlations above and Spearman correlations below the diagonal. Number of observations: 56224. Correlations with significance levels below 5% appear in bold print.

Table 3: Regressions

	Dependent variable:	
	Modified Jones DA	Dechow and Dichev DA
	(1)	(2)
Ln(Total assets)	-0.011^{**}	0.013^{***}
	(0.005)	(0.001)
Market to book	-0.00000^{***}	0.00000^{***}
	(0.00000)	(0.00000)
Return on assets	0.015^{*}	0.003^{*}
	(0.008)	(0.002)
Sales growth	-0.001	-0.0003^{*}
	(0.001)	(0.0002)
Estimator	ols	ols
Fixed effects	gvkey, fyear	gvkey, fyear
Std. errors clustered	gvkey, fyear	gvkey, fyear
Observations	56,224	56,224
R^2	0.019	0.013
Adjusted R^2	-0.163	-0.170
Note:		*p<0.1; **p<0.05; ***p<0.01