DOI: 10.1111/corg.12229

WILEY

SPECIAL ISSUE ARTICLE

The ABCs of empirical corporate (governance) research

Renée B. Adams

University of New South Wales, Sydney, NSW 2052, Australia

Correspondence
Renée Adams, University of New South Wales, Sydney, NSW 2052, Australia
Email: renee.adams@unsw.edu.au

Abstract

Manuscript Type: Commentary

Research Question/Issue: I ask how academics can do better empirical research in corporate governance and other fields.

Research Findings/Insights: I argue that academics can do better research by using old methods better. I provide a list of recommendations I label as ABCs as a reminder that one can write excellent papers if one gets the basics right. The ABCs also serve as a reminder that one must master the basics before one can credibly use new techniques.

Theoretical/Academic Implications: Following the recommendations I provide can improve research by making the research more transparent and easier to replicate.

Practitioner/Policy Implications: Better research improves policy-making and practice, especially if improvements in research practices lead to research that is easier to read.

KEYWORDS

Corporate Governance, Research, Methods, Assumptions, Causality, Writing

As scientists, we are conditioned to constantly look for new methods to advance our knowledge. But, sometimes we can achieve our goals more easily by using old methods better. This commentary lists some ABCs of empirical work that I think are useful to remember. The term "ABCs" is deliberate. It serves as a reminder that one can write excellent papers if one gets the basics right. It is also a reminder that one must master the basics before one can credibly use new techniques. As Hamermesh (2000, p. 378) says: "Before we resort to wizardry, we should be certain that we do not add confusion by making mistakes with simpler techniques."

These ABCs are based on my experience, my reading of the literature, and on my extensive experience as a referee and editor. Until it became untenable, I used to have a policy that I would referee at least once for any journal that asked me. As a result, I have refereed a substantial number of papers at more than 40 journals at all levels of impact across a variety of different disciplines (e.g. finance, economics, accounting, management, and psychology). I've also been an associate editor, special issue editor, and now editor. As a result, I've followed the progress of a substantial number of papers at the journals and know what referees complain about.

My ABCs are not exhaustive, nor are they meant to be. Doing empirical work is a craft (e.g. Hamermesh, 2000). Real world data typically does not satisfy the assumptions of the models. Often, we do not even know what the right model is, i.e. there is model uncertainty. As a result, the same problem can be approached in different ways. Compelling empirical work combines methodological tools with creativity, critical thinking, intuition, judgement, taste, and writing skills. As such, it would be difficult, and even counterproductive, to try to formalize an exhaustive list of ABCs for doing empirical work. The "ABCs" depend, in large part, on the research question. As our questions evolve, so will the ABCs.

THE ABCS

A. Assumptions: Discuss them!

Referee complaints about the validity of instrumental variables or other features of the paper's identification strategy are central to most rejections. A surprising number of these rejections could have been avoided if the authors had discussed the assumptions necessary for identification. As soon as one verbalizes the arguments why instrumental variables or treatments are exogenous, one often realizes they are not. This realization can inform the search for a better instrument or "shock" and ultimately help authors write a more compelling paper.

B. Big data: More data is not always better!

In this era of big data, we are conditioned to think more data is always better. But, it is not necessarily true. Some phenomena are better studied in the cross-section, making panel data unnecessary. For example, the question whether CEO ownership affects performance is really a cross-sectional question because ownership varies very little over time. Can it hurt to have more data than necessary? Maybe. Suppose one estimates the relation between CEO ownership and performance in panel data. Without firm fixed effects, statistical significance is overstated. With firm fixed effects, one either has problems of multicollinearity if the panel is too short (e.g. Hamermesh, 2000; Zhou, 2001) or one identifies the wrong thing—in this case the effect of CEO turnover instead of the effect of changes in incentives.[1] In either case, it is hard to conclude that CEO ownership affects performance. Thinking about the source of variation in the data and discussing it is important—not only for reducing data collection costs, but also for addressing the question at hand.

C. Causal effects or correlations? Discuss!

In Finance and Economics, most authors now strive to identify causal effects. Other disciplines are also starting to worry about identification (e.g. Antonakis, 2017; Antonakis, Bendahan, Jacquart, & Lalive, 2010). But, identification is difficult, especially in corporate governance research. Only a few papers can credibly claim to identify causal effects. Most other papers seem to fall into two camps: those that completely avoid a discussion of causality and those that misapply identification techniques, such as diff-in-diff (see D) or matching (see M), to claim to identify casual effects. Our scientific knowledge might be better advanced if authors (and referees and editors) were honest about the difficulties inherent in identification. Not discussing causality at all is out of the question. But, in the absence of perfect identification (to the extent it exists), a thoughtful discussion of what biases might exist in the results and what we can learn from documented correlations might serve our purposes better than overstating our ability to identify causal effects. Miller (2013) puts it nicely: "causal methods leverage qualitative insights for quantitative purposes, which is the only way to leap from correlation to causation." This means qualitative insights have to come first. Only then can quantitative methods follow.

D. Difference-in-difference: Does it look like a medical experiment?

Difference-in-difference is the identification technique du jour, in part because it is easily implemented. But, this ease of implementation is deceptive. It is never easy to identify causal effects. Settings that satisfy the assumptions necessary for diff-in-diff to identify causal effects are hard to find. If the setting does not remotely look like a setting in which all subjects are "sick" and the "medicine or placebo" is randomly assigned, i.e. if the situation does not resemble a medical experiment, the diff-in-diff probably does not identify causal effects.

[1]Since most CEOs have a relatively short tenure, their ownership will not vary significantly over their tenure. So, big changes in CEO ownership within a firm will arise primarily from changes in the CEO.

E. Econometric godliness is less important than data cleanliness! (Hamermesh, 2000)

The quality of empirical work is only as good as the quality of the data. Holderness (2009) provides a good example of the severe consequences of poor data quality. Holderness argues that the literature may have overemphasized the importance of agency problems between managers and shareholders. One reason is that academics too readily accepted the arguments in Berle and Means (1932), even though Berle and Means based their arguments on inaccurate data. Specifically, Berle and Means classified firms with missing data as "believed to be widely held," which meant their statistics were biased towards finding that firms were managerially controlled.

F. Fixed effects: Why not?

Fixed effects are designed to strip out means. Deliberately demeaning some variables but not others can lead estimators to be inconsistent (see Gormley & Matsa, 2014). If you are not using fixed effects when you could, referees become suspicious about the sources of variation in your data and the statistical significance of your results (see B).

G. Grammar: Use it or lose the R&R!

Research shows spelling and grammar errors negatively impact one's perceived intellect, trustworthiness and conscientiousness. If you can't be bothered to spellcheck your paper or make sure the grammar is correct, why would referees trust your claims to have clean data (see E)? Computers cannot understand spelling mistakes. If you make many spelling mistakes in the paper, how can readers trust your code?

H. Heteroskedasticity and group correlation: Correct for it!

Intuition tells us that there is no such thing as homoskedasticity in empirical corporate (governance) research. Standard errors are correlated across individuals, firms, industries, countries, and events. As a result, homoskedastic standard errors are underestimated and t-statistics overestimated (e.g. Petersen, 2009). Correcting for heteroskedasticity and group correlation (and telling the reader that you've done it) is essential.

I. Identification strategy: Discuss!

If you claim causal effects (or your writing suggests you are), discuss your identification strategy!

J. Just because something is published does not make it right!

The publishing process is not perfect. Unlike in the physical sciences, we cannot replicate a paper's results in the lab. This means referees and editors have to use their subjective judgement in evaluating papers. Moreover, standards change over time. Econometrics is an evolving field (see e.g. Athey & Imbens, 2017). Not too long ago, the concept of weak instruments did not exist. The proper use of polynomials in regression discontinuity designs is still being debated (e.g. Gelman & Imbens, 2016). It is not a good idea to use ideas/methods from published papers (even in top journals) unquestioningly—especially if they are central to the identification strategy.

K. Kinky results: What do they mean?

Check that the results are plausible and intuitive. The referees will.

L. Literature: Read it!

Simply citing what other people have said about the literature, instead of reading it oneself, is not a good idea. Like in children's Telephone or Whisper games, passing on what others say can lead to distortions. Similarly, it is bad form to neglect discussing literature that is inconsistent with your results. Informed readers (i.e. referees) will trust the results less if the literature discussion is biased.

M. Matching: Use it correctly!

Matching is often misused to address endogeneity problems. Matching assumes "unconfoundedness" or "selection-on-observables." But, one also controls for observables in regressions. In corporate governance, we typically worry about unobservables, not observables. Matching offers little improvement in causal inference relative to regression, because, in a nutshell, it cannot address omitted variable problems. Should one still use it to make regressions more effective? It is unclear, especially when the subsample of matched observations is small.

N. Number of instruments: Increasing it does not increase instrument validity!

Instrumental variable exogeneity cannot be tested (N=1)! Using multiple instruments to leverage overidentification tests of instrument exogeneity is not a shortcut. Overidentification tests are not valid without at least one exogenous instrument. The exogeneity of that instrument cannot be tested. The credible use of instrumental variables is not a matter of statistics, it is a matter of qualitative insight and persuasive writing.

O. OLS is a good friend!

Regardless of the functional form of the conditional expectation function, Angrist and Pishke (2009) argue regression may be of interest because it is the best linear approximation to the conditional expectation function. So, ordinary least squares (OLS) is a useful benchmark for more sophisticated methods.

P. Path dependence: Should you take the road less travelled by?

It is impossible to innovate on every dimension in one paper. One must build on prior literature. A useful exercise is to check whether one is doing something because everyone else is doing it or because one truly thinks it is the right thing to do.

Q. Question what you read!

The author's goal is to sell you a story. The reader's job is to make up his or her own mind.

R. Replicability enhances credibility!

Replicability is becoming more and more of a concern. While providing the final data and code for published results helps, it may not be enough. Chang and Li (2015) find that they could replicate only 59% of papers at journals requiring authors to provide data and code (see also Galiani, Gertler, & Romero, 2017). To be replicable, papers need to clearly describe the data and the assumptions necessary to transform the raw data into the final data set.

S. Significance: Not just about stars!

In the end, we care about results that are economically significant, i.e. the magnitudes of the effects are big enough to matter. In large data sets, almost everything will be statistically significant (see B), so the number of stars has little meaning.

T. Treatment on the treated or on the control?

Absence of treatment can also be a treatment. For example, when studying the effect of bank bailouts on bank behavior, it is tempting to think bailed-out banks are in the "treatment" group. But, not-being-bailed-out is also a treatment that is likely to affect the behavior of banks in the "control" group. When absence of treatment is a treatment, diff-and-diff can be misleading.

U. Understand the institutional setting!

Acquiring a good understanding of the institutional setting can be useful for developing identification strategies or avoiding potentially flawed identification strategies (see Catan & Kahan, 2016). Making sure the reader understands the institutional setting enhances credibility and ensures readers feel they have learned something (see C).

V. Visualization: A picture can convey a thousand words!

Visualization is useful if it can convey the key result simply. It is almost essential in diff-in-diff and regression discontinuity studies. If you do not provide the picture visualizing treatment effects or behavior around the cut-off, referees will wonder why.

W. Winsorization: Why?

Because OLS minimizes sums of squared residuals, it weights outliers heavily. A good discussion of the importance of outliers can be found in Guthrie, Sokolowsky, and Wan (2012). But, winsorization changes the data. Data is fundamental to empirical work (see E), so if one is going to winsorize, one should explain why. At the very least, it is a good idea to discuss what concern about outliers motivates winsorization, to say how many observations have been affected by winsorization, and to show results with and without winsorization (or include a dummy for winsorized observations in the analysis). Similar arguments can be made about trimming.

X. Xpress yourself well!

Writing is one of the most important tools in empirical work. Yet, many referees complain that they cannot understand what the author is trying to say. As McCloskey (1999) says: "Writing is thinking." As Cochrane (2005) says: "Many economists falsely think of themselves as scientists who just 'write up' research. We are not; we are primarily writers … Most good economists spend at least 50% of the time they put into any project on writing. For me, it's more like 80%."

Y. Y do you include lagged Y as a control variable?

Including lagged dependent variables as a control hard-wires an endogeneity problem that must be addressed using dynamic panel data methods, e.g. Arellano–Bond estimators. Since the instruments for Arellano–Bond techniques come from within the system, it is tempting to include lagged Y in a model to avoid having to search for exogenous instruments. But, Arellano–Bond methods are not designed to solve the endogeneity of X problem.[2] They are designed to solve the endogeneity of lagged Y problem. Unless economic reasoning suggests you need to include lagged Y, it is better not to.

Z. Zombies are for movies!

A zombie paper is a paper that pretends to be alive when it is actually dead, i.e. it was already written by someone else. Authors who fail to keep up with the literature run the risk of resurrecting zombies. Referees tend to be experts in the field. Referees may even have written the original paper. Referees like to kill zombies.

ACKNOWLEDGEMENTS

I am especially grateful to Rob Tumarkin for helpful discussions about writing. I am also grateful to Wing Wah Tham, Praveen Kumar, and Alessandro Zattoni for helpful comments.

REFERENCES

Angrist, J., & Pischke, J. S. (2009). *Mostly Harmless Econometrics* (1st ed.). Princeton University Press.

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, *28*, 5–21.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*, 1086–1120.

Athey, S., & Imbens, G. (2017). The state of applied econometrics—Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3–32.

Berle, A., & Means, G. (1932). *The modern corporation and private property*. New York: Macmillan.

Catan, E. M., & Kahan, M. (2016). The law and finance of antitakeover statutes. *Stanford Law Review*, *68*, 629–680.

Chang, A. C., & Li, P. (2015). Is economics research replicable? Sixty published papers from thirteen journals say "usually not". In *Finance and Economics Discussion Series 2015-083*. Washington, DC: Board of Governors of the Federal Reserve System. https://doi.org/doi:%2010.17016/FEDS.2015.083

Cochrane, J. (2005). Writing tips for PhD students. Retrieved from https://faculty.chicagobooth.edu/john.cochrane/research/papers/phd_paper_writing.pdf (accessed October 12, 2017).

Galiani, S., Gertler, P., & Romero, M. (2017). Incentives for replication in economics. UC Berkeley Working Paper. Retrieved from http://www.paulgertler.com/uploads/4/7/5/1/47512443/galiani-gertler-romano_replication_in_economics_nber.pdf (accessed October 12, 2017).

Gelman, A., & Imbens, G. (2016). Why high-order polynomials should not be used in regression discontinuity designs. Columbia University Working Paper. Retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/regression_discontinuity_16sep6.pdf (accessed October 12, 2017).

Gormley, T. A., & Matsa, D. A. (2014). Common errors: How to (and not to) control for unobserved heterogeneity. *The Review of Financial Studies*, *27*, 617–661.

Guthrie, K., Sokolowsky, J., & Wan, K.-M. (2012). CEO compensation and board structure revisited. *The Journal of Finance*, *67*, 1149–1168.

Hamermesh, D. S. (2000). The craft of labormetrics. *ILR Review*, *53*, 363–380.

Holderness, C. (2009). The myth of diffuse ownership in the United States. *The Review of Financial Studies*, *22*, 1377–1408.

McCloskey, D. N. (1999). Economical writing *(2nd ed.)*. Long Grove, IL: Waveland Press.

Miller, M. K. (2013). The uses and abuses of matching in political science. Working Paper, George Washington University. Retrieved from https://sites.google.com/site/mkmtwo/research (accessed October 12, 2017).

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, *22*(1), 435–480. doi:10.1093/rfs/hhn053

Zhou, X. (2001). Understanding the determination of managerial ownership and its relationship to firm performance. *Journal of Financial Economics*, *62*, 559–571.

How to cite this article: Adams R. The ABCs of empirical corporate (governance) research. *Corp Govern Int Rev*. 2017;25:461–464. https://doi.org/10.1111/corg.12229

[2]Arellano–Bond techniques use lagged values as instruments. Referees typically hate lagged values of X as instruments for X.