

E IS FOR ENDOGENEITY

Endogeneity in Content Analysis & Finding Strong Instruments to Resolve It

John R. Busenbark

*Academy of Management
2017 Content Analysis PDW
August 5th, 2017*

The University of Georgia

TERRY 

COLLEGE OF BUSINESS

Benefits & Detriments of Content Analysis

Content analysis *quantifies text* to *measure* psychological *constructs* – what people think, feel, and do – by what they say

- **Many empirical benefits**

- Psychometric advantages (Aguinis & Edwards, 2014)
- Internal validity compared to archival (Short et al., 2010)
- External validity compared to self-reports (McKenny et al., 2012)
- More proximal than archival (McKenny et al., 2016)

- **Detriments owing to endogeneity**

- Omitted variable bias (Semadeni et al., 2014)
 - Are subjects thinking something else that relates to the variable?
- Measurement error (Kennedy, 2008)
 - Does the measure remain the same across many contexts?

Endogeneity in Content Analysis

- Relationship between regulatory focus and M&A activity

$$y_{(\text{M\&A activity})} = \alpha + \beta_{1(\text{Reg.Focus})} + \beta_{2-n(\text{controls})} + e$$

Everything correlated with y
(net income) not included in
the model

Post-acquisition
integration capabilities

Maybe regulatory focus is
really M&A confidence

$$\text{Corr}[x, e] \neq 0$$

“Instead of estimating the ‘true’ relationship between the independent and the dependent variable, OLS regression **mistakenly includes** the correlation between the independent variable and the **error term**” (Semadeni et al., 2014: 1072)

The Solution: Two-Stage Models

First Stage

$$X_{(\text{Reg.Focus})} = \alpha + \beta_{1-2(\text{instruments})} + \beta_{3-n(\text{predictors})} + u$$

Second Stage

$$Y_{(\text{M\&A})} = \alpha + \beta_{1(\text{predicted } x)} + \beta_{2-n(\text{controls})} + e$$

Instruments – At least one (hopefully two) variables not correlated with “u” or “e”

- ***It ALL depends on instruments***
 - *Corrected model must be specified correctly!*
 - “Weak instruments can report results that are inferior to those reported by OLS...” (Semadeni et al., 2014: 1070)
 - No instruments “can often do **more harm than good**” (Kennedy, 2008: 271)
- ***Good instruments are hard to find these days....***
 - ***MUST*** be exogeneous (Kennedy, 2008)
 - ***MUST*** be strong predictors (Stock et al., 2002)
 - Almost ***no SMJ articles*** from 2005-2012 had good ones (Semadeni et al., 2014)
 - “Completely inappropriate” tests for good instruments (Larcker & Rusticus, 2010: 192)

A Glimmer of Hope

- **Instruments in content analysis**
 - Several content-analysis related instruments
 - Finding the right one depends on the context
 - Whether the content analysis constructs is the IV, DV, or both

- **Independent variable is from content analysis**
 - E.g., **Regulatory focus** predicts M&A activity (Gamache et al., 2015)
 - E.g., **Mentioning phrases** predicts analyst downgrades (Busenbark et al. forthcoming)
 - Number of characters in a document
 - Number of documents released
 - Number of relevant sections in the document
 - Size of the document/size of images in the document
 - Other content analysis constructs discriminant from the DV

More Glimmers of Hope

- **Dependent variable is from content analysis**
 - E.g., Firm wrongdoing on **tenor of media coverage** (Zavyalova et al., 2012)
 - E.g., CSR and **media praise** (Petrenko et al., 2016)
 - Archival financial data
 - Arellano-Bond estimation
 - Industry characteristics
 - Geographic characteristics
- **Independent and dependent variables from content analysis**
 - E.g., **Positive media** predicts more **positive media** (Pollock et al., 2008)
 - E.g., **Entrepreneurial language** predicts **legitimacy** (Gao et al., 2016)
 - Number of characters in IV documents
 - Financial data theorized to predict IV
 - Arellano-Bond estimation if outcome is to an event
 - Executive demographic characteristics

A Couple Caveats on these Instruments

- **Some can represent important constructs**
 - Number of characteristics reflects info complexity (Loughran & McDonald, 2014)
 - Number of documents can reflect litigation or abnormal behavior (Donelson et al., 2012)
 - Ensure industry or geography is not an important component of the DV construct (Wang et al., 2014)
 - Arellano-Bond estimation is requires great precision (Arellano & Bond, 1991)
- **Don't jump into two-stage models too quickly**
 - They're not the most efficient
 - May present significant Type II errors
 - Think through your content analyses constructs conceptually

Thank You!

ANY QUESTIONS?

References

- Aguinis, H. & Edwards, J. R. 2014. Methodological wishes for the next decade and how to make wishes come true. *Journal of Management Studies*, 51(1): 143-174.
- Arellano, M. & Bond, S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2): 277-297.
- Bettis, R., Gambardella, A., Helfat, C., & Mitchell, W. 2014. Quantitative empirical analysis in strategic management. *Strategic Management Journal*, 35(7): 949-953.
- Busenbark, J. R., Lange, D., & Certo, S. T. forthcoming. Foreshadowing as impression management: Illuminating the path for security analysts. *Strategic Management Journal*, doi:10.1002/smj.2659.
- Donelson, D. C., McInnis, J. M., Mergenthaler, R. D., & Yu, Y. 2012. The timeliness of bad earnings news and litigation risk. *The Accounting Review*, 87(6): 1967-1991.
- Gamache, D., McNamara, G., Mannor, M., & Johnson, R. 2015. Motivated to acquire? The impact of CEO regulatory focus on firm acquisitions. *Academy of Management Journal*, 58(4): 1261-1282.
- Gao, H., Yu, T., & Cannella, A. A. 2016. The use of public language in strategy: A multidisciplinary review and research agenda. *Journal of Management*, 42(1): 21-54.
- Hamilton, B. H. & Nickerson, J. A. 2003. Correcting for endogeneity in strategic management research. *Strategic Organization*, 1(1): 51-78.
- Kennedy, P. 2008. *A Guide to Econometrics* (2nd Edition ed.). Oxford, UK: Blackwell.
- Larcker, D. F. & Rusticus, T. O. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, 49(3): 186-205.
- Loughran, T. & McDonald, B. 2014. Measuring readability in financial disclosures. *The Journal of Finance*, 69(4): 1643-1671.
- McKenny, A. F., Short, J. C., Zachary, M. A., & Payne, G. T. 2012. Assessing espoused goals in private family firms using content analysis. *Family Business Review*, 25(3): 298-317.
- McKenny, A. F., Aguinis, H., Short, J. C., & Anglin, A. H. 2016. What doesn't get measured does exist: Improving the accuracy of computer-aided text analysis. *Journal of Management*: doi: 0149206316657594.
- Petrenko, O. V., Aime, F., Ridge, J., & Hill, A. 2015. Corporate social responsibility or CEO narcissism? CSR motivations and organizational performance. *Strategic Management Journal*, 37(2): 262-279.
- Pollock, T. G., Rindova, V. P., & Maggitti, P. G. 2008. Market watch: Information and availability cascades among the media and investors in the US IPO market. *Academy of Management Journal*, 51(2): 335-358.
- Semadeni, M., Withers, M. C., & Certo, S. T. 2014. The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, 35(7): 1070-1079.
- Short, J. C., Ketchen Jr, D. J., Combs, J. G., & Ireland, R. D. 2010. Research methods in entrepreneurship: Opportunities and challenges. *Organizational Research Methods*, 13(1): 6-15.
- Stock, J. H., Wright, J. H., & Yogo, M. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4): 518-529.
- Wang, L., Madhok, A., & Xiao Li, S. 2014. Agglomeration and clustering over the industry life cycle: Toward a dynamic model of geographic concentration. *Strategic Management Journal*, 35(7): 995-1012.
- Zavyalova, A., Pfarrer, M. D., Reger, R. K., & Shapiro, D. L. 2012. Managing the message: The effects of firm actions and industry spillovers on media coverage following wrongdoing. *Academy of Management Journal*, 55(5): 1079-1101.