

## Online Appendix: Similar Literature in Other Fields

Because the Heckman two-step method originated in the field of Economics, we examined recent papers in the *Journal of Economics and Management Strategy* (JEMS) to learn whether a Strategy journal more closely tied to Economics used the Heckman two-step method more correctly. Using the same technique to find cites of Heckman (1979) between 1995 and 2016, we were surprised to find that only four papers in JEMS use the Heckman two-step method.

To compare relative emphasis on different techniques, we used the database at Wiley Online to conduct a keyword search of the four journals analyzed above and of JEMS. The results appear in Table A1. First, as we expected, only 13% of JEMS articles that mention *endogeneity* also mention *Heckman*; the comparable figures for SMJ, ASQ, MS, and OS are 25%, 19%, 14%, and 19% respectively. Thus JEMS seems to be using the Heckman two-step method less when encountering endogeneity. Second, we find that JEMS contributors employ other instrumental-variable approaches: 68% of articles in JEMS that mention *endogeneity* also mention *instrument*, as compared to 51%, 44%, 65%, and 68% respectively for the other four journals. Finally, articles that name *endogeneity* and *Heckman* account for a smaller proportion of those that mention *instrument* and *endogeneity* in JEMS, at 18%, than in the other four journals at 49%, 44%, 22% and 27% respectively.

INSERT Table A1

In order to pinpoint the differences in techniques, we further analyzed the sub-sample of JEMS articles that mention both *instrument* and *endogeneity* for common trends. Of these 65 papers, 14 explicitly use two-stage least-squares, 26 use instrumental variables in OLS, and three use the Generalized Method of Moments with instruments. Furthermore, ten explicitly explain why instrumental-variable approaches would be inappropriate for their setting or their data. Of the papers that use instrumental-variable approaches, all but five (two of which explicitly use the technique as a robustness check) either employ a statistical technique or economic intuition to justify the use of the instruments; three of the articles also use the Heckman two-step method. This analysis demonstrates that journals focused on the economics literature within the Strategy field make less use of the Heckman two-step method and more use of other instrumental-variables techniques accompanied by statistical tests for the validity of the instruments. Because the Heckman two-step method was introduced in the Economics literature earlier and was used widely after its introduction, its weaknesses were first identified and discussed in that literature; other fields seem to have been slower to follow suit.

Recently, other fields that took up the Heckman two-step method with comparable fervor have begun to recognize the need for certain assumptions to be met. For example, Lennox,

Francis, & Wang (2011) explore use of the model in the Accounting literature. The authors make a similar argument as this paper: that use of the Heckman two-step method without a valid instrument (i.e., either with exclusion restrictions without an economic basis or without exclusion restrictions) is highly problematic. The authors then replicate a paper in the Accounting literature, Jackson, Liu, and Cecchini (2009), and explore the consequences of using the Heckman two-step method without a valid instrument. That paper, Jackson *et al.* (2009), estimates the effect of firms' depreciation methods on capital investment decisions while accounting for the endogeneity of depreciation method choice (Lennox *et al.*, 2011). Lennox and colleagues first point out why the paper's instruments are unlikely to satisfy the exclusion restrictions necessary to use the Heckman two-step method; they argue that the instruments used should also be included in the second-stage or outcome equation, and that this alternative specification would drastically alter the results (Lennox *et al.*, 2011). Replicating the paper's model with Compustat data, they show that the results of the selection model can vary widely with slight differences in the specification (Lennox *et al.*, 2011). Specifically, including these variables in the outcome equation drastically changes the coefficient of interest (Lennox *et al.*, 2011). First, both variables are found to be significant in the outcome equation and thus should clearly be included in the second stage. More importantly, both the coefficients on the choice variable and the inverse Mills ratio reverse direction or become insignificant when these variables are included in the outcome equation. The authors thus find that "a researcher could obtain quite literally any possible outcome when making fairly minor and apparently innocuous changes to the set of exclusion restrictions" (Lennox *et al.*, 2011: 589).

The inverse Mills ratio becomes insignificant when the two variables are added to the outcome equation; thus the authors also compare the results to OLS. They find that the coefficient on the choice variable in OLS coincides with their model and again contradicts the findings of Jackson *et al.* (2009) (Lennox *et al.*, 2011). Forcing an exclusion restriction by omitting variables in the outcome equation will lead to bias, but using the Heckman two-step method without exclusion restrictions is also problematic, as indicated earlier.<sup>1</sup> Thus, exclusion restrictions need to be justifiable on "economic grounds" (Lennox *et al.*, 2011: 595).

The fields of Sociology and Criminology have produced similar work. In Sociology, selection models are widespread: many Sociology studies examine the effects of social programs that will not be random in assignment (or at least not random in takeup). Berk (1983) introduced the method to the Sociology literature; his paper, like many others, indicates that the vector of exogenous variables in the selection and outcome equations "may, or may

---

<sup>1</sup>Note that both of these scenarios are investigated in the simulation analysis.

not, be the same” without explaining the consequences (Berk, 1983: 390). However, Berk (1983) does emphasize that a “theory of selection” is necessary for correct application of the model (Berk, 1983: 396).<sup>2</sup> Winship and Mare (1992) emphasize the need to correct for heteroskedasticity in the errors, although current econometric programs typically make this correction automatically (Winship & Mare, 1992). The authors refer to the bias that can occur in small samples and the inconsistency that results when selection is large. As in the work in Accounting, the authors close by suggesting alternative methods to the Heckman two-step method when these issues are highly problematic.

Bushway, Johnson, and Slocum (2007) conduct a meta-analysis similar to ours but in the field of Criminology. The authors examine 25 recent papers that use the Heckman two-step method and identify common problems, including issues with correlation and adjustment of standard errors as well as use of the Heckman two-step method without exclusion restrictions (Bushway *et al.*, 2007). As mentioned earlier, because the Heckman two-step method does not technically require exclusion restrictions, many authors do not include them, a decision that can exacerbate other issues with the method. The authors present methods to examine the extent of selection, as well as a test to investigate whether the Heckman two-step method can be used appropriately in a given setting, previously discussed in Lueng and Yu (1996) (Bushway *et al.*, 2007).

## Bibliography

Berk, R. A. (1983). “An introduction to sample selection bias in sociological data”, *American Sociological Review*, 386-398.

Bushway, S., Johnson, B. D., Slocum, L. A. (2007). “Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology”, *Journal of quantitative criminology*, 23(2): 151-178.

Jackson, S. B., Liu, X. K., Cecchini, M. (2009). “Economic consequences of firms depreciation method choice: Evidence from capital investments”, *Journal of Accounting and Economics*, 48(1): 54-68.

Lennox, C. S., Francis, J. R., Wang, Z. (2011). “Selection models in accounting research”, *The Accounting Review*, 87(2): 589-616.

Heckman, J. (1979). “Sample Selection Bias as a Specification Error”, *Econometrica*, 47(1): 153-161.

Leung, S. & Yu, S. (1996). “On the Choice Between Sample Selection and Two-Part Models”, *Journal of Econometrics*, 72: 197-229.

---

<sup>2</sup>In the conclusion, Berk mentions the importance of exclusion restrictions in reducing multi-collinearity and reducing the weight put on functional form (Berk, 1983).

Winship, C. & Mare, R.(1992) “Models for Sample Selection Bias”, *American Review of Sociology*, 18: 327-350.

**Table A1: Article Counts by Journal**

	JEMS	SMJ	ASQ	MS	OS
Article Contains "Endogeneity"	96	399	73	357	151
' and Contains "Instrument"	65	205	32	232	103
Article Contains "Endogeneity" and "Heckman"	12	100	14	50	28
' and contains "Instrument"	9	52	10	40	20
(Endogeneity + Heckman) / ( Endogeneity)	12.50%	25.06%	19.18%	14.01%	18.54%
(Endogeneity + Heckman) / (Endogeneity + Instrument)	18.46%	48.78%	43.75%	21.55%	27.18%
(Endogeneity + Instrument) / (Endogeneity)	67.71%	51.38%	43.84%	64.99%	68.21%
(Endogeneity + Heckman + Instrument) / (Endogeneity + Heckman)	75.00%	52.00%	71.43%	80.00%	71.43%

**Journal**

Journal of Economics and Management Strategy  
 Strategic Management Journal  
 Administrative Science Quarterly  
 Management Science  
 Organization Science

**Abbreviation**

JEMS  
 SMJ  
 ASQ  
 MS  
 OS