Online Appendix for

Director Histories and the Pattern of Acquisitions

Peter L. Rousseau and Caleb Stroup*

October 17, 2015

*Rousseau, Peter.L.Rousseau@Vanderbilt.Edu, Department of Economics, Vanderbilt University, Box 1819 Station B, Nashville, TN 37235 USA. Stroup, castroup@davidson.edu, Department of Economics, Davidson College, Box 7123, Davidson, NC 2803.

This appendix includes robustness tests of our hypothesis that a historical interlock increases the probability that a given directional firm pair merges. These include issues related to time-invariant (Section A) and time-varying (Section B) selection of directors, a consideration of matching estimators (Section C), robustness tests for several assumptions applied in forming our main sample (Section D), and the role of the relative size of the potential acquirer and target in determining whether acquisitions occur among historically-interlocked firm pairs (Section E).

A. Time-Invariant Director Selection

In this section we address the "spurious correlation" that would arise from director-firm matching using firm-pair specific fixed effects to control for unobservable and time-invariant features that are specific to an acquirer, target, or acquirer-target pair. These persistent common factors could include industry positioning, product lines, investment advisors, corporate governance structures, network connectedness, CEO entrenchment, board and firm size, presence of antitrust pressure, and profitability, among many others. Because the number of firm pairs increases proportionally with the sample size, constructing these fixed effects is computationally infeasible for binary response models due to the incidental variables problem in which the number of parameters increases in proportion to the number of firm pairs.[1] Fortunately, consistent estimates can be obtained by de-meaning a linear probability model, which is econometrically equivalent to including these fixed effects.

Table A.1 presents the estimated marginal effects of historical interlocks, using the full sample described in Section III.B and used to estimate equation (4). Column (1) includes

[1]Even if these effects were computable, latent variable frameworks such as the logit do not permit computation of the variance of individual effects, so the estimated coefficients are identified only up to a scale factor. This prohibits comparative estimates of how controls for pair-specific effects alter parameter values (Wooldridge (2002), 470).

only historical interlocks, column (2) adds contemporaneous interlocks to the regression, and column (3) includes the complete set of controls from the full specification shown in column (5) of Table 3. In all three equations, a historically-interlocked firms is about 6.5 times more likely to merge relative to an average firm. These estimates are larger than those obtained from our main logit models, which is expected given that the downward-biased coefficients typical of binary response models in rare-events data.

[Table A.1 here]

B. Time-Varying Director Selection

While theory suggests that time-invariant unobserved heterogeneity is the most likely form of director selection, in principle time-varying unobserved heterogeneity could be driving the results. To address this, we use the insights developed in Altonji, Elder, and Taber (2005) to gauge the degree to which time-varying director selection might be at play. The technique quantifies the amount of selection bias that would be required to explain the entire effect of director experience on acquisition patterns.[2]

Bellows and Miguel (2009) develop a general statistic that makes no assumptions on the shape of the error distribution. This statistic measures how much greater the influence of unobservables on selection must be relative to the influence of observables on selection to fully remove the estimated effect of the variable of interest. The statistic is $\theta_r = \hat{\beta}^f/(\hat{\beta}^r - \hat{\beta}^f)$, where $\hat{\beta}^r$ is the estimated coefficient from a regression with a restricted set of controls and $\hat{\beta}^f$

[2]This procedure is motivated by the insight that the amount of selection on observables conveys information about the amount of selection on unobservables. All that is required is that the amount of selection on observables be at least as large as the amount of selection on unobservables. Altonji *et al.* (2005) argue that this assumption is no less implausible than the assumptions required for OLS estimation.

is the coefficient from a regression with the full set of controls. The key choice in constructing this statistic is to select the restricted set of controls appropriately, so we work with several different restricted models. An estimated ratio of unity means that selection on unobservables must be at least as strong as selection on observable characteristics to account for the entire baseline estimate. Similarly, a number greater than one, say 3, would mean that selection on unobservables must be three times greater than selection on observables to attribute the main effect to director selection. Numbers less than one imply that negative selection is present and that the true effect of director experience is in fact larger than the baseline estimates.

Panel B of Table A.1 presents the results, again employing the sample used to estimate equation (4). The first column shows the estimate when the restricted set consists of a constant only. In this case, the statistic indicates that selection on unobservables would need to be almost four times greater than selection on observables to fully explain the effect of historical interlocks on acquisition patterns. A similar result obtains in columns (2) and (3), which respectively add year fixed effects and contemporaneous interlocks. Column (4), which adds the full vector of control variables, provides an estimated statistic of about 4.3.

The finding that firm-pair fixed effects do not cause attrition in the effect of historical interlocks along with the finding that time-varying unobserved heterogeneity is unlikely to affect empirical validity offer evidence that director selection does not play a crucial role in explaining the observed effect of historical interlocks on acquisitions.

C. Matching Estimators

We now estimate the effects of historical interlocks on merger decisions using matching methods. These nonparametric estimates control for endogeneity bias by comparing the conditional merger propensities of historically-interlocked firm pairs (the "treatment" group) with observationally similar firm pairs that are not historically-interlocked (the "control" group). In doing so, these estimates also provide an alternative approach to constructing the counterfactual set of firm pairs that are not historically interlocked.

We implement the estimators using the standard two-stage propensity score matching procedure, which uses the sample from equation (4) and estimates a probit regression of a historical interlock on the full vector of firm-pair characteristics and then uses the estimated parameters to construct predicted conditional treatment probabilities, i.e., the propensity score (Moffitt (2004)). The propensity scores are then used to form untreated firm pairs to match with each historically-interlocked firm pair. When each treated pair is matched with a single untreated pair, this procedure is known as "one nearest-neighbor matching." The estimated treatment effect is the expected difference in merger probabilities between the treated and control groups (Abadie and Imbens (2006)). For robustness and comparison, we also report unmatched treatment effects along with "two" and "three" nearest-neighbor matching estimates.

[Table A.2 here]

Table A.2 shows the estimated treatment effects. Panels A through C report unmatched estimators in column (1) along with one, two and three nearest-neighbor matching estimators in columns (2)-(4) using several sampling approaches. Coefficients represent the percent increase in acquisition probability associated with a historical interlock with t-statistics in brackets.

Panel A reports estimates obtained by implementing the matching procedure on a pooled cross-section that includes one observation for each ordered firm pair in the full sample described in Section III.B. The dependent variable takes a value of unity if firm i acquires firm j during the 1996-2006 period and zero otherwise, while the historical interlock takes a value of unity if a firm-pair is historically-interlocked at any point from 1996-2006 and zero otherwise.[3] The propensity score then uses the full set of firm-pair observables from equation (4) to match each of the 4,488 treated pairs with one, two, or three counterfactual

[3]Similarly, control variables are constructed as means across the 1996-2006 period.

firm pairs drawn from the possible set of 1,401,955 untreated pairs. The two nearest-neighbor matching estimates, for example, are constructed from a sample of 13,464 observations of which 8,976 are untreated. The estimates of the effect of a historical interlock are large and statistically significant in all cases, with a treatment effect of historical interlocks that increases the probability firm i acquires firm j by a factor of about 15 for three nearest-neighbor matching, which is much larger than the increase implied by the logit estimates in our main analysis (Table 3). This is once again expected given that binary response models lead to downward-biased coefficients in rare-events data.

Panel B reports estimates obtained by implementing the propensity score matching procedure on a cross-section of observations from the year 2004, drawn from the full sample described in Section III.B. The unit of observation is an ordered firm pair, the dependent variable takes a value of unity if firm i acquires firm j and zero otherwise. Here, the matching estimates are based on a sample of 1,212 treated observations that are matched based on propensity score with one, two, and three nearest-neighbors from our untreated set. The two nearest-neighbor matching estimates, for example, are constructed from a sample of 3,636 observations of which 2,424 are untreated. Here again the estimated treatment effects of historical interlocks are positive, large, and statistically significant.

Panel C restricts the full sample to the subset of firm pairs where both firm i and firm j fall into the same 4-digit SIC code. Here, the set of 905 treated firm pairs is matched with one, two, and three nearest-neighbor subsamples drawn from a universe of untreated firm pairs within 4-digit SIC codes. The estimated treatment effects are again statistically significant with magnitudes similar to those obtained with the pooled cross-section.

These large effects of historical interlocks suggest that neither the number of observations nor assumptions made about the counterfactual set of untreated pairs are driving our main results.

D. Further Robustness Checks

We now report results from several additional robustness checks. Our historical interlocks exclude instances where one of the acquirer's current directors served on the potential target's board within the past two years. We did this to avoid cases where the historically-interlocked director served on both boards during the merger planning process.

To check the sensitivity of our results to this assumption, we construct two alternative measures of the historical interlock. In the first, the historical interlock takes a value of unity for a potential acquirer in year t if at least one of its current directors served on the potential target's board in the past, but not in the past year. In the second, the historical interlock takes a value of unity for a potential acquirer in year t if at least one of its current directors served on the potential target's board in the past, but not in the past three years.

[**Table A.3 here**]

Columns (1) and (2) of Table A.3 re-estimate the main equation from column (5) of Table 3 by replacing the standard historical interlock with the one-year cutoff (column (1)) and the three-year cutoff (column (2)). The estimated coefficient on a historical interlock is larger in column (1) and smaller in column (2) than in our main results. These findings apply more generally: in additional regressions (not shown) the estimated coefficient on a historical interlock continues to decline gradually as we eliminate observations with more recent board service at the target, possibly reflecting the fact that information becomes outdated as time passes.[4]

[4] We also examined whether the effect of a historical interlock increases with the interlocked director's tenure at the potential target prior to moving to the acquirer, and found there was no statistically signifcant difference between the impact of directors with more or less experience at the target.

We next check whether our main result is sensitive to a more restrictive threshold for the percent of the target bought by the acquirer. Column (3) of Table A.3 presents re-estimates of the main equation from Table 3, but in this case excludes partial-share acquisitions, i.e., we code $ACQ_{ijt} = 1$ only if the purchasing firm acquires 100% of the target firm's shares in the deal.[5] The findings are virtually identical to our main results.

We also examine whether our main result is sensitive to inclusion of partial acquisitions. We do this by again re-estimating equation (4) retaining only complete acquisitions, i.e., we allow partial share purchases but require that the acquirer own all of the target's shares as a consequence of the deal.[6] This, for example, includes mergers where the acquirer previously held 20% of the target's shares and bought 80% through the deal. The results, presented in Column (4), are nearly identical to our main findings.

We next check to see whether our inclusion of small-value acquisitions affects our main finding. To do this, we include in the main regression only announced deals with a money value greater than 100 million dollars.[7] These findings are presented in Column (5) with an estimated effect of historical interlocks nearly identical to that in Table 3. Finally, in Column (6), we check the robustness of our results to allowing small-share acquisitions that are at least 10% of the target's shares. As before, the estimated coefficient on the historical interlock is quite similar to the main coefficient in Table 3.

E. Exploring the Role Played by Relative Size of Acquirer and Potential Target

It is also at least plausible that directors tend to move from smaller to larger firms over their careers given that larger firms tend to have higher director compensationand directors might on average choose larger and more prestigious boards as they gain experience. If true,

[5] This procedure excludes 71 announced mergers from the sample.

[6] This restriction eliminates 66 deals from the sample.

[7] This restriction eliminates 54 deals from the sample.

the firms most likely to be acquirers (i.e., larger firms) would also be those most likely to have historical interlocks with smaller targets (i.e., smaller firms).

Although we already control for the ratio of the assets of the potential acquirer and target in all of our main specifications, we report further robustness checks for the validity of the alternate story in Table A.4. We begin by checking for nonlinear effects of relative firm size in Column (1), which re-estimates the full specification from Table 3 and includes in addition second- and third-degree polynomials of relative size. These non-linear controls are not statistically significant, and their inclusion does not impact the estimated effect of historical interlocks.

[Table A.4 here]

Column (2) re-estimates the main equation on the subsample of matched firm pairs for which the ratio of acquirer and target size is in the interval between 0.75 and 1.25, allowing us to focus on the effect of historical interlocks for similarly-sized firm pairs. The estimated effect of historical interlocks on the probability of acquisition is 6.2 times greater for similarly-sized firms, consistent with the matching estimates presented in Table A.3.

Column (3) re-estimates the main equation from Table 3 and omits the control for relative size. This allows us to examine the relative importance of the size and information channels. The estimated effect of historical interlocks is extremely similar to the main estimate from Table 3, with a historically-interlocked firm pair being about 4.4 times more likely to merge relative to a pair that is not historically interlocked.

References

Abadie, A., and G. W. Imbens. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74 (2006), 235–267.

Altonji, J. G.; T. E. Elder; and C. R. Taber. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic schools." *Journal of Political Economy*, 113 (2005), 151–184.

Bellows, J., and E. Miguel. "War and Collective Action in Sierra Leone." *Journal of Public Economics*, 11-12 (2009), 1144–1157.

Moffitt, R. A. "Introduction to the Symposium on the Econometrics of Matching." *Review of Economics and Statistics*, 86 (2004), 1–3.

Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press (2002).

Note to Table A.1

Panel A reports estimates from linear probability models for the likelihood of a merger expressed as the percent increase in the probability that acquirer i announces an acquisition of potential target j in period t. Column (1) includes the historical interlock and year effects only. Column (2) includes contemporaneous interlocks and column (3) includes the full vector of controls from Table 3. All estimating equations include fixed effects for years and for ordered firm pairs. T-statistics derived from robust standard errors clustered at the firm-pair level appear in parentheses beneath the coefficient estimates. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Each cell of Panel B reports ratios based on the coefficient for historical interlocks from regressions on the probability that firm i acquires firm j in year t. The ratio is calculated $\beta^f/(\beta^r - \beta^f)$, where β^r is the estimate on historical interlocks obtained from an equation using a restricted set of controls and β^f is the estimate on historical interlocks from the full regression which includes firm-pair specific fixed effects. The restricted sets of controls are a constant (column (1)), year fixed effects (column (2)), year fixed effects and contemporaneous interlocks (column (3)), and year effects, contemporaneous interlocks, and the full set of controls less firm-pair fixed effects (column (4)).

Note to Table A.2:

The table reports matching estimates for the pair-specific match propensities. The first column reports unmatched estimates of the average treatment effect. Columns (2), (3) and (4) report one, two and three nearest-neighbor matching estimates of the effect of historical interlocks on acquisitions. Panels A-C report matching estimates obtained from various sampling methods (described in Section C of this appendix). Average treatment effects are percentage increases relative to the control group. T-statistics appear in parentheses. *, ** and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Note to Table A.3:

The table reports estimates from logit regressions for the pair-specific match propensity where the dependent variable is unity if firm i announced an acquisition of potential target j in period t. The variable of interest is the historical interlock (defined in Section III.A). All regressions include the full set of controls used in column 5 of Table 3. Columns (1) and (2) replace the historical interlock with analagous measures (defined in Section D of this appendix) that require a minimal number of years (one and three, respectively), since the historically-interlocked director last served on the target's board. Column (3) excludes partial-share acquisitions from the sample. Column (4) includes partial share acquisitions but requires that the acquirer would own all of the target's shares after the deal. Column (5) requires that deal value be equal to or greater than 100 million U.S. dollars Column (6) requires that the acquirer obtain at least 10 percent of the target as a result of the deal. Robust standard errors clustered at the firm-pair level appear in parentheses beneath the coefficient estimates. *,**, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Note to Table A.4:

The table reports estimates from logit regressions for the likelihood of a merger where the dependent variable is unity if firm i announced an acquisition of potential target j in period t. The estimating equations have the same form as the final column of Table 3 in the main paper. Column (1) includes polynomials of relative firm size. Column (2) restricts the sample to firm pairs for which the potential acquirer and target's relative size lies in the interval (1.25, 0.75). Column (3) omits controls for relative size. All regressions include fixed effects for years. Robust standard errors clustered at the firm-pair level appear in parentheses beneath the coefficient estimates. *,**, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.1
Linear Fixed Effect Models and
Estimates of Director Selection

{Insert Table A.1 Note Here}

Panel A: Effect of Historical Interlock from Fixed Effects Model

Variable	(1)	(2)	(3)
Historical interlock	643.4***	652.4***	652.3***
	(4.48)	(4.48)	(4.48)
Contemporaneous interlock		26.0***	26.0***
		(2.39)	(2.39)
Full vector of controls	no	no	yes
Year fixed effects	yes	yes	yes
Firm-pair fixed effects	yes	yes	yes

Panel B: Extent of Unobserved Heterogeneity

Variable	(1)	(2)	(3)	(4)
θ_r	3.98	3.98	4.28	4.28
Year fixed effects	no	yes	yes	yes
Contemporaneous interlock	no	no	yes	yes
Vector of controls	no	no	no	yes
Firm-pair fixed effects	yes	yes	yes	yes

Table A.2
Matching Estimators

{Insert Table A.2 Note Here}

	Unmatched	1 Neighbor	2 Neighbor	3 Neighbor
	(1)	(2)	(3)	(4)
Panel A: Pooled Cross-Section				
Treatment Effect	2,501	2,071	2,069	1,551
t-statistic	[31.62]	[9.73]	[10.14]	[10.00]
Observations	1,401,955	8,976	13,464	17,952
Panel B: Annual Cross-Section (2004)				
Treatment Effect	1,790.5	700.1	1,500.1	2,299.9
t-statistic	[33.44]	[2.34]	[2.62]	[2.70]
Observations	792,257	2,424	3,636	4,848
Panel C: Within 4-Digit Industry Sample				
Treatment Effect	5,435.8	2,700.0	1300.0	1580.0
t-statistic	[36.94]	[5.09]	[4.90]	[5.00]
Observations	507,996	1,810	2,715	3,620

Table A.3
Further Robustness Tests

{Insert Table A.3 Note Here}

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Historical interlock	4.006***	2.406***	3.857***	3.875***	3.890***	3.844***
	(0.184)	(0.429)	(0.195)	(0.193)	(0.191)	(0.190)
Contemporaneous	1.609***	1.447***	1.019***	1.155***	1.205***	1.129***
interlock	(0.248)	(0.247)	(0.283)	(0.268)	(0.269)	(0.266)
Relative size	0.004***	0.004***	0.003***	0.003***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Relative sales to assets	0.004	0.004	-0.060	-0.053	0.005	0.005
	(0.007)	(0.007)	(0.037)	(0.033)	(0.007)	(0.007)
Relative market to book	0.033***	0.042***	0.032***	0.033***	0.026***	0.042***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Same 4-digit SIC	3.070***	3.128***	3.140***	3.110***	3.036***	3.128***
	(0.111)	(0.111)	(0.119)	(0.118)	(0.117)	(0.114)
Same county	1.107***	1.274***	1.010***	1.086***	1.085***	1.123***
	(0.157)	(0.151)	(0.170)	(0.166)	(0.166)	(0.161)
Year fixed effects	yes	yes	yes	yes	yes	yes
Industry fixed effects	yes	yes	yes	yes	yes	yes
Observations (thousands)	6,951	6,951	6,951	6,951	6,951	6,951
Pseudo R^2	0.142	0.122	0.144	0.147	0.141	0.146

Table A.4
Effects of a Historical Interlock on Merger Likelihood

{Insert Table A.4 Note Here}

Variable	(1)	(2)	(3)
Historical interlock	3.854***	4.119***	3.810***
	(0.189)	(0.561)	(0.189)
Contemporaneous interlock	1.186***	1.194**	1.158***
	(0.261)	(0.577)	(0.260)
Relative size	0.010***	1.371	
	(0.003)	(1.005)	
Relative size squared	-0.000		
	(0.000)		
Relative size cubed	0.000		
	(0.000)		
Relative sales to assets	0.004	-0.032	0.005
	(0.007)	(0.060)	(0.007)
Relative market to book	0.031***	-0.008	0.033***
	(0.006)	(0.042)	(0.005)
Same 4-digit SIC	3.077***	2.770***	3.056***
	(0.112)	(0.344)	(0.114)
Same county	1.143***	2.219***	1.143***
	(0.158)	(0.411)	(0.159)
Year fixed effects	yes	yes	yes
Industry fixed effects	yes	yes	yes
Observations (thousands)	6,951	1,121	6,951
Pseudo R^2	0.14	0.20	0.14