

Clougherty, Joseph; Duso, Tomaso

Working Paper

Correcting for self-selection based endogeneity in management research: A review and empirical demonstration

DIW Discussion Papers, No. 1465

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Clougherty, Joseph; Duso, Tomaso (2015) : Correcting for self-selection based endogeneity in management research: A review and empirical demonstration, DIW Discussion Papers, No. 1465, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/108982>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2015

Correcting for Self-Selection Based Endogeneity in Management Research A Review and Empirical Demonstration

Joseph Clougherty and Tomaso Duso

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2015

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

Correcting for Self-Selection Based Endogeneity in Management Research: A Review and Empirical Demonstration

Joseph Clougherty¹

University of Illinois at Urbana-Champaign and CEPR
350 Wohlers Hall; 1206 S. 6th St., MC-706; Champaign, IL 61820; USA
Tel: +1 217 333 6129
E-mail: jaclough@illinois.edu

Tomaso Duso

DIW Berlin and DICE, Heinrich-Heine University
Mohrenstrasse 58, 10117 Berlin; Germany
Tel: +49 30-89789-520
E-mail: tduso@diw.de

Abstract

Foundational to the discipline of management is the idea that organizational decisions are a function of expected outcomes; hence, the customary empirical approach to employ multivariate techniques that regress performance outcome variables on discrete measures of organizational choices (e.g., investments, trainings, strategies and other managerial decision variables) potentially suffer from self-selection based endogeneity bias. Selection-effects represent an internal validity threat as they can lead to biased parameters that render erroneous empirical results and incorrect conclusions with regard to the veracity of theoretical assertions. Our review of the empirical literature suggests that the issue of selection-effects has received increasing attention in management; yet, the techniques to correct for selection-effects have not always been employed in the proper manner, thus estimations often suffer from design shortcomings that potentially render flawed empirical findings. We explain the nature of self-selection based endogeneity bias and review the techniques available to researchers in management to correct for selection-effects when organizational decisions are discrete in nature. Employing data on M&A investment decisions and rival-firm value reactions, we provide empirical examples that demonstrate the tradeoffs involved with the alternative techniques.

Keywords: Endogenous treatment, event-study, Heckman, mergers, rival firms, Switching regressions model

JEL classification: M1, M2, C1

¹ Corresponding author

Introduction

Fundamental to management research is the ambition to make causal claims; yet, our discipline infrequently allows for the randomized experiments which represent the gold standard for making strong causal inferences (Antonakis et al., 2010). While Bloom et al. (2013, 2014) represent rare examples where such experiments have taken place, Bascle (2008) points out that randomized controlled experiments in the managerial context are often infeasible due to reasons of ethics, expense, and unwillingness of managers and businesses to be randomly placed into treatment and control groups. Accordingly, management researchers customarily rely on observational data sets and regression models where independent variables cannot be exogenously manipulated (Li, 2012). Such empirical contexts yield far greater potential for endogeneity bias to manifest.

Endogeneity bias renders coefficient estimates from standard regressions causally uninterpretable as the estimates will be inconsistent in the sense that they do not converge to true coefficient values; i.e., the estimates derived when the sample size approaches the census of activity. Wooldridge (2002) outlines and clarifies the three sources of endogeneity bias: measurement error, simultaneity, and omitted variables. Measurement error in variable constructs can both attenuate and bias the effect of regression estimators, while simultaneity occurs when one of the predictors is jointly determined along with the dependent variable (Li, 2012). Yet as Bascle (2008) points out, omitted variables have received the greatest amount of attention by management scholars as the principal source of endogeneity. Omitted-variable bias arises when an omitted – or latent – factor exists which both affects the dependent variable and is correlated with one or more explanatory variables. In essence, such a condition ensures that included regressors will correlate with the error term – as variation in the latent variable will manifest in the error term – and this violates the most important of the OLS assumptions (the exogeneity assumption): i.e., that the error term has an expected value of zero given any explanatory variable (Wooldridge, 2013).

While Heckman (1976, 1979) first observed that neglecting selection represents a specification error that is akin to the omitted-variable bias, the relationship between selection bias and endogeneity bias has not always been well appreciated by management scholars (Echambadi et al., 2006). The basic insight behind selection bias being a form of omitted-variable bias is that the selection process

represents an excluded variable that manifests in the error term and correlates with the endogenous choice construct and the outcome variable (Antonakis et al., 2010). Heckman's (1976, 1979) seminal contribution involves modeling the selection process as a truncation problem so as to introduce a variable – the inverse Mills ratio – into the substantive equation of interest in order to correct for the selection bias. In essence, this variable – which captures and corrects for the selection process – is missing when standard regression techniques are employed. Moreover, Antonakis et al. (2010) make clear that omitted selection is an important source of endogeneity bias.

Selection-based endogeneity manifests in two main forms: sample-selection and self-selection biases. Heckman's (1976, 1979) foundational work was principally motivated by sample-selection problems, as samples can be non-representative of a true population and thus threaten both internal and external validity—see Berk (1983) for an excellent review of this issue. Yet Heckman (1979) was also conscious of the analogous self-selection problem (our focus here), as he observed that comparing the wages of management trainees with the wages of non-trainees can result in biased estimates of treatment effects. In the self-selection context, bias derives not from sample selection (i.e., no bias exists in the scope of the sample being studied) but instead the studied agents make choices regarding assignment into the mutually-exclusive treatment and non-treatment groups based on unobservables that correlate with both outcomes and observable predictors. For example, worker traits that are unobservable to the researcher (e.g., diligence, innate intelligence, etc.) may determine both the selection into management trainee programs and future wages, thus a parameter estimating the relationship between participation in a trainee program and the future wages of a manager can be confounded with the selection process into that trainee program. In fact, Wooldridge (2002) deems self-selection to be a common source of omitted variable bias in empirical work being done in the behavioral and social sciences.

Self-selection based endogeneity clearly represents an internal validity threat to research throughout management, as the decisions, processes, trainings, investment types, strategies and other discrete business phenomena which we study will be chosen (i.e., selected into) by managers with outcomes in mind. Accordingly, the exact nature of self-selection bias and the methods which are employable to correct for this form of endogeneity will be the focus of this article. Yet our

contribution does not represent the first work to point out that the management discipline should be mindful of the dangers involved for researchers when organizational decisions are self-selected. Shaver (1998) first decisively observed that firms choose strategies based on their attributes and implications, therefore organizational choices are endogenous and self-selected. Shaver cautions then that the customary approach in management to regress performance measures on strategy choice variables could lead to misspecification and incorrect positive and normative conclusions by failing to account for self-selection effects. Shaver prescribed the Heckman (1979) procedure as an effective approach to deal with issues of self-selection in organizational decisions.

Influenced by Shaver's (1998) seminal study, Hamilton and Nickerson (2003) make clear that the field of strategic management is particularly subject to self-selection based endogeneity, as the field's fundamental premise is that managerial decisions regarding strategy choice are endogenous to their expected performance implications. In other words, managers do not make organizational decisions – such as appropriate organizational form, use of markets or hierarchies, and type of investment undertaken – on a random basis. Instead, the outcomes – or performance implications – of these decisions are central to the ultimate organizational decision taken. Omitted variables are then likely to affect both the strategy choice and the performance outcome, thus rendering biased the coefficient estimates from standard regressions of strategy type on performance outcomes. In essence, both Shaver (1998) and Hamilton and Nickerson (2003) point out an inherent contradiction involved with the fact that empirical work in strategic management seldom controlled for self-selection effects; namely, the existence of the field is predicated on the idea that managerial decisions are endogenous to expected outcomes (i.e., organizational decisions are not best characterized as randomized processes), hence self-selection effects should actually be both endemic and fundamental to empirical studies of discrete strategies.

We want to underscore that the potential for self-selection bias resides in empirical work throughout the management discipline. Yet the influential articles by Shaver (1998) and Hamilton and Nickerson (2003) specifically targeted the field of strategic management in outlining the nature of selection-based endogeneity and providing prescriptions with respect to correctly addressing this issue by modeling the selection process. For instance, Hamilton and Nickerson's (2003) survey of the

Strategic Management Journal (SMJ) – which arguably best represents strategy research in general – for the years 1990 through 2001 delivered some sobering findings: only 27 out of 196 empirical papers on firm performance actually correct for endogeneity issues; and only 3 of these papers used some type of instrumental-variable technique (Bascle, 2008). Accordingly, if correct procedures to deal with self-selection based endogeneity are to have penetrated the management discipline, then one might expect the strategy field to best exhibit the adoption of such methodological approaches in response to these notable urgings. In other words, strategic management should seemingly be at the frontier within management when it comes to correctly addressing self-selection based endogeneity.

We take off from the Hamilton and Nickerson (2003) review by surveying SMJ for the years 2002 through 2014 in order to gather whether variants of the Heckman (1979) technique have been employed by strategy scholars to deal with the evident selection-effects in this field. We find that 47 studies in SMJ attempt to correct for potential sample-selection effects (Appendix A), while 55 studies attempt to correct for potential self-selection effects over this period (Appendix B). Accordingly, the issue of selection-based endogeneity has clearly received increasing attention in the period following the Shaver (1998) and Hamilton and Nickerson (2003) studies. Yet while a number of empirical studies have adopted methods to correct for selection bias, these studies are often characterized by design defects that potentially render uninformative empirical results. Our review of the literature suggests that many studies fail to properly identify the selection equation as constructs that are unique to the selection equation are often not specified; i.e., 8 of the 55 self-selection studies clearly do not involve identification variables in the selection equation, while an additional 13 studies potentially do not involve identification variables in the selection equation. Yet even when identification variables are specified, the rationale behind the identification assumption (i.e., why the variables impact selection but not the main equation of interest) is oftentimes not set out by the authors. This inability to confidently conclude accurate identification of selection effects is compounded by the fact that many studies do not report the results of the Heckman procedure or of the selection equation; instead, such procedures tend to be auxiliary tests that are simply reported to editors and referees and are accordingly repressed from final publication.

Thus, despite the efforts to prod strategy scholars to adopt the appropriate methods to deal with selection-based endogeneity, it is fair to say that empirical work in the field still does not consistently employ the correct methodological practices. The correct procedures to deal with selection effects do not appear to have deeply penetrated the management literature in that many of the empirical studies addressing selection effects do so in a manner that falls short of best practice. Moreover, it is reasonable to assume then that the deficiencies in the strategic management field are indicative of deficiencies with regard to dealing with selection effects in the greater management literature. Antonakis et al. (2012) make clear the dangers involved with such practices as violating essential design and methodological conditions might mean that yielded coefficient estimates cannot be interpreted as even indicating correlation nevertheless causation.

As to what might explain the inconsistent and inadequate attempts by empirical scholarship in the management literature to correct for selection-based endogeneity issues, a few potential causes arise. First, the prescriptive literature in management has advocated different techniques – e.g., Shaver (1998) proffers the Heckman procedure, Hamilton and Nickerson (2003) proffer Lee's (1978, 1982) switching-regression procedure, and Basile (2008) argues that an IV approach can handle many self-selection issues – which potentially leaves researchers unsure as to the appropriate method(s) to employ. Second, the prescriptive literature in management arguably has not provided sufficient and clear information on the inevitable upsides and downsides involved with the alternative methods to deal with selection bias. Accordingly, the literature within management dealing with issues of selection bias could gain from an integrative study that helps researchers understand the correct approach to take in order to deal with selection-effects, as making the appropriate decision with regard to correcting for selection bias depends on the available data and the precise research question being faced.

We aim to fill the above gap in the literature concerning methodological practices by providing empirically-minded researchers in management clear guidelines with regard to choosing between the different methodological techniques and correctly estimating these procedures for dealing with selection-based endogeneity. Accordingly, our modern contribution attempts to provide value for empirical researchers in management facing the potential issue of self-selection bias as well as the

related issue of sample-selection bias. In order to deliver on these aims, we first discuss in detail the nature of selection-based endogeneity problems. Furthermore, we will outline and review the different methods – with the attendant advantages and disadvantages – available for researchers to deal with self-selection bias. We will provide some practical advice for estimating these procedures so as to ensure that best practices are employed and that the methods are correctly used. We also present an empirical demonstration of the self-selection issue in the context of a binary organizational decision. In particular, we employ data on mutually-exclusive M&A decisions and the rival-firm value reactions to these mergers in order to demonstrate the tradeoffs involved with the alternative techniques to deal with self-selection bias.

The Self-Selection Based Endogeneity Problem

Before delving further into the issue of self-selection bias, it is important to provide further grounding on the related issue of sample-selection bias, as this sets the basic logic for more complex models of selection. Sample-selection bias can be present when the researcher uses non-randomly selected samples to estimate causal relationships. The issue often arises in practice for two reasons: when the observational units make decisions so that a subset of a particular population is not observed; and when samples of observational data involve some selection by analysts and data processors (Heckman, 1979). Accordingly, empirical situations where the outcome variable is only observed for a portion of the true sample often yield sample-selection issues as a censoring in the data is present. For instance, Heckman (1974) observes – in his motivational application piece – that estimating the return on education for women involves sample-selection issues and biased coefficient estimates because the researcher only observes outcomes – i.e., wages – for working women. Yet the decision to work (i.e., to select into employment and thus select into a sample) is surely endogenous and possibly driven by some of the same omitted drivers of an individual's wage. Accordingly, the estimated return on education in such a selected sample would be biased and unrepresentative of the average return on education in the entire population of women. It should be reiterated that for the selection to generate biased coefficient estimates it is crucial that non-randomness exists in the data-construction process. Thus by excluding some observations from a population of activity in a systematic manner, the

estimates that are actually inferred from the biased sample might actually be a product of random perturbations (Berk, 1983).

Self-selection bias is of a different nature than sample-selection bias, as in these empirical contexts there are no issues with regard to the dependent variable not being observed for relevant subsamples of the population. Instead, the self-selection concern arises when the dependent variables are observed for different subsamples, yet a non-randomness is involved with the manifestation of these dependent constructs. Staying with the illustrative context of worker wages, Lee (1978) studies the impact of unionism on wages where he has data on the wages of workers who are either unionized or non-unionized; hence, sample-selection bias is not at play. Yet workers make decisions as to whether they join a union or not; moreover, this decision is endogenous in the sense that unobserved factors (e.g., intelligence, connections, etc.) can affect that decision to join the union, but such unobserved factors can also affect future wages. Lee (1979: 977) summarizes the above when he states that in such cases “decisions are based on the possible outcomes under alternative choices and observed outcomes are final outcomes of the decision process. So decisions and outcomes are interrelated”. Yet if standard regression techniques were to be employed in such an empirical context, the error terms would violate the necessary assumption of an expected mean of zero.

This self-selection problem boils down, in essence, to the problem of treatments not being randomly assigned to the agents being studied in observational data sets (Antonakis et al., 2010). Instead of being randomly assigned into treatment and non-treatment groups, the organizational choices we study in management are a function of managers attempting to enhance outcomes. Since treatments are not randomly assigned to groups of agents, untreated observational agents do not necessarily represent adequate counterfactuals with respect to treated observational agents. The root of the problem is that observational agents (e.g., managers, firms, transactions) experiencing the treatment (e.g., engaging in a particular strategy) might substantially differ from those agents not experiencing the treatment in terms of observable and unobservable factors. It is, of course, the unobserved characteristics which can yield pernicious effects in terms of biases as they are subsumed in the error term. Under such conditions, it is quite possible that the residual in the selection equation into the treatment will correlate with the residual in the main equation of substantive interest.

Therefore, the treatment (e.g., a particular strategy choice) will be correlated with the residual term in the main equation. In this sense, selection-based endogeneity represents a sub-set of the omitted-variable problem, as omitted variables in the residual will correlate with the treatment—a violation of the assumption residing behind standard regression techniques.

We can provide a grounded example of how omitting selection can be problematic and lead to biases when self-selection is present. Let us consider, for example, the overall performance implications (e.g., the ROA) for firms when deciding to enter new market niches. King and Tucci (2002) engage in such a study and are fundamentally interested in whether entry into new product areas positively or negatively affects a firm's ROA (i.e., performance). Yet firm entry is a choice variable by managers (i.e., not randomly assigned by the researcher), thus it is quite possible that unobservable factors (e.g., the intangible 'dynamic capabilities' of firms) exist which determine whether a firm chooses to enter a new market segment or not. Thus, the entry decision is explained by other factors that are not observed in the substantive equation of interest. In essence, an additional 'selection equation' explains the product entry decision of firms; i.e., that binary decision by organizations to obtain treatment. If, however, those latent factors (i.e., those unobserved dynamic capabilities) are common across the substantive equation and the selection equation, then the greater self-selection problem is that in such an empirical context the error terms of the two equations will be correlated, thus the entry decision will correlate with the error term in the substantive equation which violates the strict exogeneity assumption.

Accordingly, latent sources of variation (i.e., omitted variables) which manifest in the error term and are correlated with the treatment represent direct validity threats to causal inferences. Yet in this context, the omitted factor is not just some simple omitted construct – as is custom with general endogeneity concerns – which happens to correlate with the dependent and independent variables, but instead it is the selection process itself which is omitted from the model. Going back to our use of counterfactuals in order to provide an intuitive grounding of self-selection bias: the fundamental issue here is that without controlling for the selection process, one cannot use firms that decide not to enter a new market niche as adequate counterfactuals in order to estimate the performance impact of those firms which do decide to enter a new market niche.

Endogeneity based on self-selection has the potential to undermine a great deal of the empirical literature residing within the management discipline. Since much of what we study involves how organizational decisions impact performance outcomes of one sort or another, the selection into these organizational decisions – and the attendant potential for biases in coefficient estimates – is intrinsic to our research. Furthermore, selection-based endogeneity does not affect coefficient estimates in a consistent manner, as both inflation and deflation of treatment effects can materialize and even the sign of coefficients can be reversed (Berk, 1983). Accordingly, researchers (e.g., Modrek and Cullen, 2013) cannot make the claim that their hypothesized results hold despite the presence of uncorrected selection. It is also important to underscore that not only will the coefficient estimates of variables subject to selection be inconsistent in standard regressions, but there are also potential negative spillover effects to other right-hand-side constructs. In particular, regressors that correlate with variables subject to selection-based endogeneity will also find their coefficient estimates being adjusted to the extent that they correlate with the problematic variables (Antonakis et al., 2010). The greater danger is that researchers neglect self-selection as a fundamental source of endogeneity and that by doing so biased parameters result which leads to erroneous and inconsistent findings and ultimately incorrect conclusions with regard to the veracity of certain theories (Hamilton and Nickerson, 2003).

Alternative Methods to Deal with Selection-based Endogeneity

Only by properly modelling the selection process can a researcher correctly deal with selection-based endogeneity issues. The choice of the correct modelling procedure, however, requires a clear understanding of both the source and the specific nature of the endogeneity problem. Accordingly, we set out here the different sources of selection-based endogeneity bias and match these sources with appropriate methodological procedures that might be employed to deal with such selection effects. The different methodological procedures are based on Heckman's (1974, 1976, 1979) seminal contribution – which lead to his 2000 Nobel Prize in Economics – yet it is important to underscore that these different selection models vary both in terms of motivation and implementation. We first consider the

issue of sample-selection bias (and the attendant Heckman procedure to deal with this bias) before turning to the more nuanced issue of self-selection based endogeneity.

Sample selection

As previously mentioned, the sample-selection problem reduces to a censoring issue, as the outcome variable is only observed for a subsample of the population. Heckman's (1974) seminal work considered this issue as he pointed out that sample selection can be viewed as a sort of omitted-variable problem. Heckman's basic idea to deal with this issue was to first employ instruments that predict selection into a treated group (i.e., estimate a proper selection equation) and then via that procedure create a variable that captures the differences between the sample for which we observe the outcome and the sample for which we do not observe the outcome. In a second step, he proposes to introduce into the substantive equation of interest a hazard rate – the so-called inverse Mill's ratio – which is a monotone decreasing function of the probability that a particular observation for the variable of interest is observed. This allows the researcher to control for selection-based variance—variance that would otherwise manifest in the error term. By engaging in such a two-step procedure, a researcher can obtain unbiased coefficient estimates when sample-selection issues are present. The covariance matrix estimated by OLS in the second stage is, however, inconsistent; hence, crucial to making correct causal inferences with such two-step procedures is the need to consider the appropriate estimation of the variance-covariance matrix—a topic we will discuss extensively. Empirical researchers often neglect, unfortunately, the issue of what is the appropriate estimation of the variance-covariance matrix, thus leading to potentially misleading inferences.

Let us now formalize Heckman's argument which will help us more-precisely discuss the appropriate econometric approach to this issue. Researchers are often interested in investigating how the exogenous and observable characteristics (X) of a particular agent i affect an outcome variable Y . Thus, they often aim to estimate the following model:

$$Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon_i = X_i' \alpha + \varepsilon_i. \quad (1)$$

In cases where the data is censored, the outcome Y_i is only observed in one subsample of the population, $z_i = 1$. Heckman proposes to model this process via a latent variable approach which

assumes that the outcome of interest is observed if and only if an unobserved latent variable (z_i^*) exceeds a particular threshold:

$$z_i = \begin{cases} 1 & \text{iff } z_i^* > 0 \\ 0 & \text{otherwise} \end{cases}. \quad (2)$$

Moreover, the latent variable can be expressed as a linear function of the observed (Z_i) and unobserved (ω_i) characteristics of agent i :

$$z_i^* = \beta_0 + \beta_1 Z_i + \omega_i = Z_i' \boldsymbol{\beta} + \omega_i.$$

To estimate this model, the researcher has to make an (untestable) distributional assumption regarding the error term (ω_i). Generally, it is assumed that the error term is normally distributed with mean 0 and variance σ_ω^2 , which is normalized to one for the sake of identification. Thus, equation (2) becomes a probit model and $\Pr(z_i = 1) = \Phi(Z_i' \boldsymbol{\beta})$ where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.² If the outcome variable is only observed in the subsample $z_i = 1$, then the full model consists of equation (1) and (2).

In such an empirical context, the sample-selection problem is due to ε being correlated with ω , thus X is correlated with ε in violation of the strict exogeneity assumption. Specifically, if the variance-covariance matrix of the full model is represented as follows:

$$\Omega = \begin{pmatrix} \sigma_\varepsilon^2 & \rho \sigma_\varepsilon \sigma_\omega \\ \rho \sigma_\varepsilon \sigma_\omega & \sigma_\omega^2 \end{pmatrix},$$

where σ_ε^2 and σ_ω^2 are the variances of the error terms (and the latter is normalized to one) and ρ is the correlation coefficient between the two error terms; then $\rho \neq 0$ implies that OLS estimation of model (1) is inconsistent. Indeed, latent sources of variation (i.e., omitted variables) which manifest in ε and are correlated with X represent direct validity threats to causal inferences. Yet in this context, the omitted factor is not just some simple omitted construct – as is custom with general endogeneity concerns – but instead the selection process itself is actually omitted from the model.

² It is also assumed that the error term follows a logistic distribution which generates the logit model.

To formalize this reasoning, consider the OLS regression of Y on X and take the expectation of model (1):

$$E(Y_i|X_i, Z_i = 1) = E(X_i'\alpha + \varepsilon_i|X_i, Z_i^* > 0) = X_i'\alpha + E(\varepsilon_i|\omega_i > -Z_i'\beta) = X_i'\alpha + \rho\sigma_\varepsilon \left(\frac{\phi(Z_i'\beta)}{1-\Phi(Z_i'\beta)} \right),$$

where $\phi(\cdot)$ is the standard normal density function. If the correlation between the error terms, ε_{it} and ω_{it} , is not zero (i.e., $\rho \neq 0$), then the selection into subsample $z_i = 1$ is not random and the OLS regression of Y on X would lead to biased coefficient estimates. In particular, the selection bias is equal to $\rho\sigma_\varepsilon \left(\frac{\phi(Z_i'\beta)}{1-\Phi(Z_i'\beta)} \right)$: the product of the covariance between the two error terms ($\rho\sigma_\varepsilon$), and a term measuring the likelihood that the observation is in the subsample $z_i = 1$. The latter term represents the inverse Mill's ratio; thus, consistent and unbiased estimation of the effect of X on Y requires taking this omitted factor into account. Yet in the case where sample-selection is random – i.e., $\rho = 0$ – then an OLS estimation will deliver consistent estimates of the parameters α .

It is also important to underscore that identification of the selection equation is of fundamental importance. As noted in our review of the literature, a number of empirical researchers employ Heckman-type procedures that do not involve unique identifiers in the selection equation. Such estimations are possible because even when the set of explanatory variables in equations (1) and (2) are identical (i.e., $X_i = Z_i$), all of the parameters (including ρ and σ_ε^2) can be identified via the functional assumptions regarding the bivariate distribution of the error terms. Yet, Maddala (1983) observes that this type of identification has proven to be quite poor and non-transparent. It is therefore essential that researchers employ additional exclusion restrictions (i.e., additional instruments) to better identify such models and, specifically, the selection mechanism.

In terms of estimation method, the researcher can choose between a Limited Information Maximum Likelihood (LIML) estimator and a Full Information Maximum Likelihood (FIML) estimator. First, by assuming that ε_i and ω_i follow a bivariate normal distribution, the researcher can estimate all of the parameters of interest via FIML by minimizing the following log-likelihood function:

$$\ln L = \sum_{z_i=0} \ln(1 - \Phi(Z_i'\beta)) + \sum_{z_i=0} \ln\left(\frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}}\right) + \sum_{z_i=1} \frac{1}{2\sigma_\omega^2} (Y_i - X_i'\alpha)^2 + \sum_{z_i=1} \ln\Phi\left(\frac{Z_i'\beta + \rho\left(\frac{Y_i - X_i'\alpha}{\sigma_\varepsilon}\right)}{\sqrt{(1-\rho^2)}}\right). \quad (3)$$

Under this assumption, the FIML estimator is consistent and asymptotically efficient; though, it might deliver inconsistent estimates if the (untestable) assumption of bivariate normality does not hold.

The LIML estimation procedure is computationally less complex and more robust as compared to the FIML estimator, but LIML is less efficient if the null hypothesis of joint normality is valid. The less severe (but still untestable) assumption that the researcher needs to employ for this estimation procedure is that the random-error term in the selection equation, ω_i , is normally distributed. The LIML estimator from this model is the original two-step estimator proposed by Heckman (1974) and consists of estimating the probit model and computing the inverse Mill's ratio. After doing so, the researcher estimates equation (1) via OLS with the inverse Mill's ratio added as an additional regressor; and, finally, the researcher should estimate a consistent variance-covariance matrix for equation (1) in order to correct for the fact that the inverse Mill's ratio is a predicted explanatory variable.³ This last step is critical and is often omitted by researchers—see, for example, the multivariate approach proffered by Hamilton and Nickerson (2003).

While a clear tradeoff between efficiency and robustness is present when considering the choice between FIML and LIML, it is difficult to prescribe when the researcher should employ which estimator. If the sample is very small, then it might be advisable to focus on the LIML estimator, though this could lead to imprecise inferences as the results might indicate an insignificant relationship when the estimator is simply too inefficient. Hence, the most consistent and transparent approach would seemingly be to estimate and report both the FIML and LIML results.

The estimation of the Heckman model is, nowadays, quite easy to implement with common statistical packages. For instance, Stata allows readily estimating the Heckman model by both FIML and LIML procedures via the 'heckman' command. This routine allows the researcher to specify several options with regard to the parameter estimates, and with regard to consistent estimations of the variance-covariance matrix in the presence of heteroskedasticity and clustering.

³ A consistent variance-covariance matrix which allows correctly measuring standard errors and other relevant statistics (t-statistics and the creation of confidence intervals) and allows making correct inferences can be estimated by using an asymptotic approximation (as in the ML case) or by bootstrapping.

Self-selection: Endogenous Treatment and Endogenous Switching

The self-selection issue – and the related endogenous-treatment and endogenous-switching problems – differs from the sample-selection issue outlined above. Furthermore, these different problems involve some subtle differences amongst each other—especially with respect to the way the treatment impacts the outcome. While the appropriate procedures to deal with all of these issues build on Heckman’s fundamental contribution and the logic we discussed above concerning sample-selection, the nature of these concerns goes beyond sample-selection as the focal issues are no longer the mere censoring of data. In self-selection cases, the outcome variable is observed for an entire sample which is assumably representative of the population.⁴ However, a potentially endogenous ‘treatment’ exists which partitions the sample population into two – or more – subsamples. Moreover, we can distinguish between two self-selection variants that we define as ‘endogenous treatment’ and ‘endogenous switching’. The main difference between the two is whether the researcher thinks that the treatment merely has an intercept effect on the outcome (as in endogenous treatment), or whether this effect is also on the coefficient estimates (as in endogenous switching).

Endogenous Treatment

Suppose that we are still interested in the aforementioned relationship between agent i ’s exogenous and observable characteristics (X) and the outcome variable Y . However, a sub-group of the full sample is affected by a treatment which is measured by the dummy variable $z_i = 1$. This model appears to be very similar to the sample-selection model discussed above; though, an important difference is that we now observe Y_i not only when $z_i = 1$, but also when $z_i = 0$. Accordingly, we can pool the entire sample and include the treatment dummy as an explanatory variable in the main equation of interest. This model can be written as follows:

$$Y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 z_i + \varepsilon_i. \quad (4)$$

⁴ Clearly, there might contemporaneously be problems of sample selection and endogeneity; however, we will not discuss such issues here.

Since the entire sample is employed, sample-selection issues are not necessarily present here. Yet this model restricts the effects of the exogenous factors X (the α 's) to be equal for both sub-groups. Thus, the effect of the treatment is to simply shift the regression line up or down (an intercept effect) conditional on the other explanatory factors. The coefficient estimate for the focal dummy variable accordingly measures the average treatment effect (ATE); i.e., the average difference of the potential outcome of the treated group as compared to the potential outcome of the group that has not been treated. In the case of this treatment being randomly assigned, the researcher could consistently and efficiently estimate this equation via OLS. But if the treatment was to be endogenous (i.e., the dummy z_{it} is correlated with the error term, ε_{it}), then such an estimation would be inappropriate and lead to inconsistent estimates. In many instances, this type of endogeneity problem can be viewed as a self-selection issue: where the observed agent endogenously chooses to be part of the treated group. As already noted, this process is related to omitted-variable bias as the fundamental source of endogeneity is that there are factors that the researcher cannot observe that affect both the self-selection process and the ultimate outcome variable.

This endogenous-treatment issue can thus be viewed as a classical endogeneity problem that can be dealt with by employing a standard instrumental variable (IV) framework. The main issue to confront in this case, however, is that the potentially endogenous variable (i.e., the dummy variable z_i above) is not continuous but is instead discrete. If the researcher is willing to make an additional assumption and take a linear approximation in order to describe the process behind the variable z_i – what Wooldridge (2013) refers to as a linear probability model – then a two-stage least squares (2SLS) estimator could be adopted in such situations. The main advantage of using the linear probability model is that the estimated coefficients in the first stage become easily interpretable. In essence, Basile (2008) submits that an IV approach can be employed in many different contexts where self-selection based endogeneity is present.

A number of disadvantages exist, however, with regard to employing an IV approach. For one, several observers have noted that the linear probability model (i.e., least-squares regressions with a dichotomous or, more generally, a qualitative dependent variable) leads to coefficient estimates that do not have known distributional properties. Hence, standard statistical inferences – such as hypothesis

testing and the construction of confidence intervals – are no longer justified in such a context. Moreover, these coefficient estimates are sensitive to the observation range of the employed data and may accordingly understate or overstate the magnitude of the true effects; in addition, they may systematically yield probability predictions that reside outside the unity range (Wooldridge, 2013). In general, the most-prudent prescriptive advice that we can provide to researchers is that they be cautious when adopting an instrumental-variable approach to dealing with self-selection issues characterized by discrete outcome variables. It is then comforting to note that our review of the empirical literature in strategic management indicates that these types of procedures have been infrequently employed: only 5 of the 55 self-selection studies in Appendix B employ an IV approach.

The most appropriate means to deal with this self-selection – or endogenous treatment – issue is then to adopt a latent variable approach that is similar to the Heckman procedure discussed for treating sample-selection issues. Hence,

$$z_i = \begin{cases} 1 & \text{iff } z_i^* > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

and z_i^* is modeled as follow:

$$z_i^* = \beta_0 + \beta_1 Z_i + \omega_i = Z_i' \boldsymbol{\beta} + \omega_i .$$

In such a context, it is again assumed that the variance-covariance matrix for model (4)-(5) is the following:⁵

$$\Omega = \begin{pmatrix} \sigma_\varepsilon^2 & \rho\sigma_\varepsilon\sigma_\omega \\ \rho\sigma_\varepsilon\sigma_\omega & \sigma_\omega^2 \end{pmatrix}.$$

It is easy to see how closely related the two models (one for sample-selection and the other for endogenous treatment in the context of self-selection) are, which makes it perhaps understandable that many researchers mix-up these two Heckman variants.

Similar to the case of sample-selection, this model can be estimated by means of both FIML and LIML estimators depending on the assumptions the researcher is willing to make with regard to the distribution of the joint error terms ε_i and ω_i . The FIML estimator requires joint normality and,

⁵ In order to yield identification, it is again assumed that $\sigma_\omega^2 = 1$.

under this hypothesis, is consistent and asymptotically efficient. The LIML estimator is consistent even if the joint-normality hypothesis fails, but LIML is inefficient if the null hypothesis of joint normality is valid. We should also note that estimations of this type can be quite easily performed with Stata by invoking the command ‘etregress’.⁶

Endogenous Switching

If the researcher thinks that the effect of the treatment is not merely a shift of the intercept but also involves differences in the relevant coefficient estimates (i.e., some or all of the α ’s differ according to the treatment), then an endogenous-switching model is necessary to employ. In essence, the endogenous switching model resembles a ‘double’ sample-selection process and can be represented as follows:

$$\begin{cases} Y_i^1 = \alpha_0^1 + \alpha_1^1 X_i^1 + \varepsilon_i^1 & \text{if } z_i = 1 \\ Y_i^0 = \alpha_0^0 + \alpha_1^0 X_i^0 + \varepsilon_i^0 & \text{if } z_i = 0. \end{cases} \quad (6a)$$

$$(6b)$$

The dummy variable (z_i) is again modeled as the dichotomous counterpart of a latent variable z_i^* and it holds that:

$$z_i = \begin{cases} 1 & \text{iff } z_i^* > 0 \\ 0 & \text{otherwise} \end{cases}. \quad (7)$$

The variance-covariance matrix is, in this model, more complex than the previous matrices and can be represented by:

$$\Omega^s = \begin{pmatrix} \sigma_{\varepsilon^0}^2 & \rho^{01} \sigma_{\varepsilon^0} \sigma_{\varepsilon^1} & \rho^0 \sigma_{\varepsilon^0} \sigma_{\omega} \\ & \sigma_{\varepsilon^1}^2 & \rho^1 \sigma_{\varepsilon^1} \sigma_{\omega} \\ & & \sigma_{\omega}^2 \end{pmatrix}.$$

The maximum likelihood (ML) estimates for this model can be shown to be consistent and asymptotically efficient; yet, the estimation may still be cumbersome because the likelihood function

⁶ A comprehensive description of the model, how it works, and how it is estimated via the command ‘etregress’ can be found in the Stata page under: <http://www.stata.com/manuals13/teetregress.pdf>.

to minimize in this case is relatively complex.⁷ Compared to the previous models, this switching-regressions model is certainly more flexible as it allows coefficient estimates to vary across sub-groups and thus estimate heterogeneous effects for each of the exogenous factors X on the outcome. These results are generally reported as the average treatment effects on the treated (ATET); i.e., the average gain from treatment for those agents who were actually treated. It is worth noting that significant correlations between the error terms (ρ^0 and ρ^1) represent the real source of endogenous selection. As an aside, simply sub-sampling – or considering the interactions between the covariates and the treatment – does not solve the fundamental endogenous-selection problem, as doing so would still mean neglecting the correlation between the two error terms.

As with the previous models, the switching-regressions model can be estimated via FIML and LIML depending on the assumptions regarding the distribution of the error terms. In particular, researchers willing to assume a tri-variate normal distribution can employ FIML to simultaneously estimate equations (7), (6a) and (6b), as such an approach would be both consistent and efficient (Lee, 1979). In order to estimate a switching-regression model via LIML and FIML, one can do so with Stata by employing the ‘`movestay`’ command.⁸

The switching-regression model can be alternatively viewed as a sort of ‘double sample selection’ problem, as the researcher could separately estimate the model for the treated and non-treated observations by means of two Heckman sample-selection procedures.⁹ In this case, the researcher would first focus on the observed data from one subsample and estimate the models (6a) and (7); and then focus on the data from the other subsample and estimate the models (6b) and (7). While taking such an approach is feasible, the fact that only variation from within the subsample

⁷ For the sake of identification, it is again normally assumed that $\sigma_\omega^2 = 1$.

⁸ Alternatively, the researcher could use the ‘`etregress`’ command where, instead of only using the treatment dummy as a regressor she also adds as additional regressors the interaction among each of the exogenous variables included in X and the treatment. This model is similar, though not identical, to (6). For a discussion see the section “Average treatment effect on the treated (ATET)” in the following document: <http://www.stata.com/manuals13/teetregress.pdf>.

⁹ As discussed above in the sub-section on sample selection, the Stata command ‘`heckman`’ can be used to perform this estimation.

would be employed in the estimations indicates that it would be a less-efficient procedure as compared to estimating the entire model in a simultaneous manner. Furthermore, undertaking two Heckman sample-selection procedures essentially means setting the parameter ρ^{01} equal zero. It is, however, worth mentioning the possibility of this type of estimation, as taking such an approach might represent a good approximation when a researcher needs to estimate a more complex structure where the selection process is multivariate and not bivariate (i.e., dichotomous) in nature but is not ordered.

Extension: Endogenous Switching with Multivariate Selection

Within the management literature that takes into account potential selection effects, the almost universal approach has been to simply consider bivariate contexts. Thus, researchers have rarely corrected for selection effects in a context where managers choose between multiple alternatives. In fact, a number of researchers have excluded additional organizational choices so that they can employ the standard bivariate procedures to deal with selection effects. For example, Brouthers et al. (2003) consider the joint-venture and wholly-owned-subsidiary entry modes but omit licensing agreements in part so they can estimate selection effects in a bivariate setting. Furthermore, the few studies which do entertain multivariate choices – or strategies – on the part of managers do not fully control for selection effects. For instance, Morrow et al. (2007) compare the stock-market reactions to three different types of strategic investments – new product introductions, mergers and strategic alliances – but instead of modeling the selection into each particular organizational decision, they instead take a bivariate short-cut by simply modelling whether the firm selects into any of the three investment types. The above indicate our literature's inability to adequately deal with situations that involve more than two organizational decisions.

The lack of studies which fully embrace an organizational context where managers make choices amongst multiple discrete variables is, nevertheless, understandable. As Berk (1983) points out, things do get quite complicated in a multivariate setting. Furthermore, Hamilton and Nickerson (2003) offer up an approach to deal with multivariate organizational decisions; however, that procedure fails to correct for the heteroskasticity involved with embedding a predicted value (the

inverse Mill's ratio) into the substantive equation of interest. Therefore, standard errors in this equation are not consistently estimated and inferences based on these standard errors are incorrect. As an aside, bootstrapping the standard errors in the second stage may represent a fix to this issue. Nevertheless, omitting multivariate contexts from our sphere of study is problematic in the sense that organizational realities often involve decision contexts that go beyond dichotomous. In short, the real world of managers involves choosing between multiple options when it comes to organizational decisions and the performance outcomes of these decisions are, of course, salient to these decisions.

Accordingly, here we go beyond a bivariate setting for organizational decisions to consider a situation where managers choose amongst multiple alternatives that are not ordinal in nature. Consider, for instance, an empirical context where the treatment goes beyond a simple partition of the sample into two sub-groups to an empirical context where the treatment partitions the sample into multiple sub-groups. For the sake of exposition, imagine that there are three possible outcomes for the treatment that managers self-select into; thus, rendering a more complex version of the model discussed above. We now have three different main equations and three sets of coefficient estimates that we must estimate, one for each subsample:

$$\begin{cases} Y_i^0 = \alpha_0^0 + \alpha_1^0 X_i^0 + \varepsilon_i^0 & \text{if } z_i = 0 \\ Y_i^1 = \alpha_0^1 + \alpha_1^1 X_i^1 + \varepsilon_i^1 & \text{if } z_i = 1 \\ Y_i^2 = \alpha_0^2 + \alpha_1^2 X_i^2 + \varepsilon_i^2 & \text{if } z_i = 2 \end{cases} \quad (8)$$

The most complex part of this empirical estimation revolves around how the researcher would model the selection process into the treatment; i.e., the selection equation for z_i . The researcher must first consider whether the outcome can be ordered or not. Second, the researcher must make an assumption with respect to the distribution of the error terms. Finally, the researcher must make some assumptions regarding an even more complex variance-covariance matrix, which can be expressed as follows:¹⁰

¹⁰ Notice that, as above, σ_ω^2 is assumed to be equal to 1 for the sake of identification. Moreover, ρ^{01} , ρ^{02} , and ρ^{12} are not defined, as Y_i^0 , Y_i^1 , and Y_i^2 are never simultaneously observed.

$$\Omega^M = \begin{pmatrix} \sigma_{\epsilon^0}^2 & \rho^{01}\sigma_{\epsilon^0}\sigma_{\epsilon^1} & \rho^{02}\sigma_{\epsilon^0}\sigma_{\epsilon^2} & \rho^0\sigma_{\epsilon^0}\sigma_{\omega} \\ & \sigma_{\epsilon^1}^2 & \rho^{12}\sigma_{\epsilon^1}\sigma_{\epsilon^2} & \rho^1\sigma_{\epsilon^1}\sigma_{\omega} \\ & & \sigma_{\epsilon^2}^2 & \rho^2\sigma_{\epsilon^2}\sigma_{\omega} \\ & & & \sigma_{\omega}^2 \end{pmatrix}$$

Without going into overdue detail, we submit that treating such an empirical context as a ‘triple sample selection’ problem represents one possible means to deal with self-selection effects in a multivariate setting. Thus, the researcher can separately estimate the model for each of the three organizational decisions. For purposes of brevity, we will not exhibit a multivariate organizational choice setting in our empirical demonstration, though we will consider a bivariate choice context where an analogous ‘double sample selection’ problems is analyzed.

Empirical Demonstration

In order to demonstrate the application of the outlined approaches to dealing with self-selection based endogeneity, we investigate how different merger types affect the underlining value of non-merging rival firms. Non-merging rival firms represent outsider firms to the merger transaction, as they compete with merging firms in the same product markets for the same customers. Moreover, the impact of a merger on the ultimate profitability and competitiveness of rival firms has been considered to be an alternative measurement procedure that allows detecting the underlining rationales and dominant effects concerning a particular merger (e.g., Chatterjee, 1992; Hitt et al., 2001; Clougherty & Duso, 2011). We measure the value effect of the merger transaction on rival firms via the cumulative abnormal return (CAR) experienced by rival firm i facing merger j taking place in year t , CAR_{ijt} .¹¹ We control for the relevant characteristics of rival and acquiring firms as these are expected to affect the CARs of rival firms. Yet, our principal explanatory variable of interest is merger type, as this represents the treatment variable that managers are likely to self-select into.

We consider a simple dichotomous strategy regarding merger type that differentiates between *efficiency-based* and *collusion-based* mergers. Hence, we partition the sample into two mutually

¹¹ We add a yearly time dimension as our data covers several years. Moreover, we need to choose control variables which are derived from balance sheet data and, therefore, are measured on a yearly basis.

exclusive categories: first, if merger j taking place in year t is efficiency-based (i.e., the portfolio of rival firms for the focal merger experience an overall value loss), then $z_{jt} = 1$; second, if the merger is collusion-based (i.e., the portfolio of rival firms for the focal merger experience an overall value gain), then $z_{jt} = 0$. The basic insight behind this categorization of mergers is that transactions which enhance collusion (or market-power) will increase the underlining value and profitability of competitor firms on average, while mergers which enhance efficiencies (or true synergies) will decrease the underlining value and profitability of competitor firms on average (Duso et al., 2007, 2011, 2013).

We chose the above empirical setting to demonstrate the alternative methodological approaches to dealing with selection bias, because self-selection based endogeneity naturally arises in the substantive regression equation. The most effective means to elicit merger type is via the effect of the transaction on the portfolio of rival firms; hence, the definition of merger type (a first dependent variable from the selection equation) implicitly depends on the effect of the transaction on a focal rival-firm's value (a second dependent variable in the substantive equation of interest). This automatically implies that selection into merger type is per-construction correlated with unobservable characteristics (i.e., the error term) that shape the effect of merger type on rival-firm values. Thus, our empirical context requires that we directly face the issue of self-selection with respect to the dichotomous merger-type variable since this regressor is likely to be endogenous and the resulting coefficient estimates biased if selection effects are not properly modelled and controlled for.

Data

Our sample of mergers is drawn from those transactions automatically requiring – due to passing a size threshold – notification to the European Commission (EC). The fundamental advantage of this data resides in the accurate identification of rival firms undertaken by the EC—an identification which derives from the initial steps taken by officials to identify the relevant product and geographic markets for each and every notified merger. The merging and rival firms were identified by carefully analyzing

and sampling the publicly available files for each transaction.¹² Some cases were dismissed due to missing stock-market information for the merging and/or rival firms. Our final sample contains information on 112 merger transactions taking place between 1991 and 2007, and 431 rival firms associated with these mergers. The size of the sample as well as the careful selection of the involved firms conforms to the best practices on the use of event studies in organizational research (e.g., McWilliams and Siegel 1997).

Our main dependent variable – the focal rival firms’ cumulative abnormal returns (CARs) – is constructed by employing the standard event-study technique. We follow standard practice in the literature by using the first day in which rumors about a particular merger appear in the press as the announcement date —identified by means of ‘Dow Jones Interactive’. Further, Thomson Financial’s ‘Datastream’ yielded information – for each rival firm i in our sample – on stock returns (SR_{id}) and market values (MV_{id}) for the two years up to day d , as well as information on the relevant market return of the industry sector (SR_{md}) where m refers to the industry. We use the Scholes and Williams (1977) method which takes returns over a 240-day period – starting 300 trading days prior to the announcement date – in order to estimate the ‘normal’ return for firm i at day d (in year t). The abnormal return is then defined as the difference between the observed return and the counterfactual return that the market model predicts; i.e., $AR_{id} = SR_{id} - \widehat{SR}_{id} = SR_{id} - \hat{\alpha}_i - \hat{\beta}_i SR_{md}$. To account for possible information leakages that can influence firm returns, we calculate the ‘cumulative abnormal return’ (CAR) as the sum of the daily abnormal returns within an event-window spanning from d_1 days before to d_2 days after the initial announcement of the merger ($CAR_i = \sum_{d=d_1}^{d_2} AR_{id}$). We use Lubatkin et al.’s (1997) relatively long short-term event window (50,5) to measure CARs.¹³ The 56-day window follows research (e.g., Seth 1990, Turk 1992, Mahoney and Mahoney 1993, Capron and Shen 2007) that stresses the need to account for pre-announcement information leakages in cases of large transactions, as well as research supporting the necessity of longer windows in order to

¹² The reports are available to the public at <http://europa.eu.int/comm/competition/mergers/cases>.

¹³ We also employed a longer window (50,50) – available upon request – to ensure robustness.

impound the impact of a merger onto the stock prices of rival firms.¹⁴ While broader windows enhance the risk of confounding events, we carefully exclude all firms which have been closely involved in simultaneous merger activity or other major events.

We also gathered information on the observable characteristics for all the involved firms in order to create a set of control factors. We matched all firms in our database with observations from Thomson Worldscope: which contains balance-sheet information for major world-wide enterprises. In particular, we generated four sets – rival, acquirer, merger and industry level variables – of variables that are in line with Villalonga and McGahan (2005) in order to make sounder causal inferences. First, our rival characteristics include: the rival's net income over current assets (Rival ROA), the rival's market to book value (Rival Tobin's Q), rival's cash equivalents over current assets (Rival Cash to Assets), the rival's long term debt over total assets (Rival Debt to Total Assets), the rival's market capitalization (Rival Market Cap), the number of two-digit SIC industries in which the rival is present (Rival's Diversification), the rivals' market value over acquirer's market value (Rival Relative Size to Acquirer). Second, our acquirer characteristics include: the acquirer's net income over current assets (Acquirer ROA), the acquirer's market to book value (Acquirer Tobin's Q), the acquirer's cash equivalents over current assets (Acquirer Cash to Current Assets), acquirer's long term debt over total assets (Acquirer Debt to Total Asset), the acquirer's market capitalization (Acquirer's Market Capitalization), and the number of two-digit SIC industries in which the acquirer is present (Acquirer Diversification). Third, our merger characteristics include: whether merging firms are based in different nations (Merger Cross-Border), and the relative size of the target with respect to the acquirer in terms of market value (Merger Relative Size Tar/Acq). Moreover, we generate a measure of the relative overlap between the merging firms in terms of lines of business (Merger Relatedness), as we do not have data that allow measuring relatedness at the resource level (e.g., Robins and Wiersema, 1995). Fourth, our industry characteristics include merger-wave drivers: a variable measuring the

¹⁴ Duso et al. (2010) confirm this reasoning when they compare – for merging and rival firms – the CARs based on different windows with measures of long-term post-merger profitability based on balance-sheet data. Especially for rival firms, they report a strong and positive correlation between CARs based on the larger (50,5) window with these profitability measures, while short-window CARs show lower or even negative correlations.

extent of industry-specific liberalization shocks (Liberalization), the industry-median sales growth (Sales growth), the industry-median return on shares (Share return), and the industry-median Tobin's Q (Tobin's q), as well as the number of notified mergers to the EC in the last two quarters (Notifications), and the square term of this construct (Notifications squared).

Our main explanatory variable – i.e., the merger type – required the generation of a 'cumulative average abnormal return' ($CAAR_j$) – i.e., a weighted average of the CARs – of the rival firms for each merger j : where the firms' individual market value (MV_i) is used to weight appropriately. Hence, the $CAAR$ for the rival firms involved in merger j are defined as $CAAR_j = \sum_{i=1}^{I_j} (CAR_i \times MV_i) / \sum_{i=1}^{I_j} (MV_i)$ where I_j is the number of rival firms involved in merger j . With these measures at hand, we follow Duso et al. (2007, 2011, 2013) in operationalizing our merger typology. Specifically, we use the $CAAR$ s of the portfolio of rival firms in order to create two subsamples of merger types: efficiency-based mergers are such that $CAAR_j < 0$, and collusion-based mergers are such that $CAAR_j > 0$.

Endogenous Treatment

If we believe merger type simply affects the intercept in our main regression equation – i.e., merger type simply shifts the CARs either up or down – then the model we would estimate is the following:

$$CAR_{ijt} = \alpha_0 + \alpha_1 R_{it-1} + \alpha_2 A_{jt-1} + \alpha_3 z_{jt} + \varepsilon_{ijt}. \quad (9)$$

Beside the average treatment effect (α_3) of the merger type (z_{jt}) on rival-firm value, we are also interested in controlling for – and understanding – the role of a large set of factors which may be important drivers of rival-firm returns and are grouped into two broad categories: rival characteristics (R_{it-1}) and acquirer characteristics (A_{jt-1}). We lag all of these variables by one year to the pre-merger period in order to reduce simultaneity concerns and elicit exogenous – or at least predetermined –

information with respect to rival and acquirer characteristics. Finally, ε_{ijt} is a random term which is assumed to be heteroskedastic, thus, we estimate White heteroskedasticity-robust standard errors.¹⁵

The merger-selection process follows a standard latent variable approach (e.g., Maddala, 1983) where each merger type is realized if, and only if, a second unobserved latent variable exceeds a particular threshold:

$$z_{jt} = \begin{cases} 1 & \text{iff } z_{jt}^* > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (10)$$

where:

$$z_{jt}^* = \beta_0 + \beta_1 R_{it-1} + \beta_2 A_{jt-1} + \beta_3 M_{jt} + \beta_4 I_{jt-1} + \omega_{ijt}.$$

With a normally distributed error term (ω_{ijt}), equation (10) becomes a probit model. We hold that the variance-covariance matrix for the bivariate process $(\varepsilon_{ijt}, \omega_{ijt})$ is characterized by a correlation coefficient ρ and variances σ_ε^2 and σ_ω^2 respectively; plus, we set $\sigma_\omega^2 = 1$ for the sake of identification. If $\rho \neq 0$, then the merger type is endogenous to rival-firm outcomes.

As discussed above, all parameters in equation (9) and (10) – including ρ and σ_ε^2 – can be identified by functional form restrictions on the distribution of the bivariate process $(\varepsilon_{ijt}, \omega_{ijt})$ by imposing joint normality. Since this identification has been proven to be poor and non-transparent, we employ additional exclusion restrictions (i.e., instruments) in order to better identify our model and, specifically, the mechanism of the selection into merger type. In particular, we assume that the selection equation is not only affected by all those factors that explain the CARs of rival firms (i.e., the characteristics of rivals and acquirers, R_{it-1} and A_{jt-1} , respectively); but also by some time-varying

¹⁵ Since the observations involving the rival firms for a particular merger would seemingly be correlated, we also clustered the standard errors at the merger level. While we obtained results for the OLS and LIML estimators that did not differ significantly from those reported, this approach did not work for the FIML estimators. This is probably due our having a large number of clusters (112) given the limited number of observations, and due to the fact that several clusters (28% of the observations) are relatively small as they involve less than four observations. For the sake of comparability among the different estimators, we always report results based on the estimation of a heteroskedasticity-robust variance-covariance matrix.

characteristics of merger j (M_{jt}), as well as characteristics of the industry where merger j takes place (I_{jt-1}). Hamilton and Nickerson (2003) also advise that industry-level characteristics represent good candidates for choosing instrumental variables when firm strategies represent the focal treatment variable.

It is, of course, important that researchers clearly outline their identification strategy and make clear the assumptions residing behind these strategies. The rationale behind our identification is that merger-specific characteristics (i.e., the geographical provenience, relative size, and horizontal product market overlap of the acquirer and target) and industry-specific characteristics (capturing the extent of the merger wave) should explain the realization of merger types, but not directly determine – or to a much lesser degree – the outcomes of specific rival-firm CARs. By invoking industry-specific characteristics, we follow Andrade et al. (2001) and Harford (2005) by constructing a set of variables which have been found to be important drivers of merger waves. Accordingly, the above instrumental variables explain the realization of merger types, but do not affect the specific CARs of rival firms; thus, conditional on merger type, these instruments are exogenous to rival-firm outcomes.

We first estimate this endogenous-treatment model by OLS and implicitly assume that merger type is an exogenous construct. This could be termed a naïve approach where obvious selection issues are neglected; however, such an approach acts as a good benchmark upon which we can compare the results where the endogenous treatment is specifically taken into account. In order to produce consistent estimates of the average treatment effect for merger type, we then estimate the model while correcting for selection effects via three different means. Namely, we employ a linear IV approach (2SLS), and then estimate the model via the two available (LIML and FIML) non-linear maximum likelihood procedures.

Endogenous Treatment Results

Table 1 reports the empirical results for the four different specifications – OLS, IV-OLS, LIML non-linear, and FIML non-linear – concerning endogenous treatment. All of these estimations include the dummy variable that captures an efficiency-based merger as the principal right-hand-side construct of

interest. The reciprocal dummy variable capturing collusion-based merger activity is repressed from these estimations as it would be unidentifiable in regression models involving a constant term since the two merger types represent mutually exclusive categories. The coefficient estimate for efficiency-based merger activity represents the differential (intercept) effect of efficiency-based merger activity as compared to collusion-based merger activity; i.e., the average treatment effect (ATE) involved with efficiency-based mergers.

We start by discussing the naïve approach where the treatment, i.e., merger type, is assumed to be exogenous; hence, the model is estimated via OLS. Amongst the set of control variables, only the coefficient estimates for Rival ROA and the Acquirer Cash to Assets ratio are significantly different from zero. Most importantly, the coefficient estimate for efficiency-based merger activity – our principal variable of interest – is not only insignificant but also has the opposite sign (a positive effect on rival firms) with respect to *a priori* expectations vis-à-vis collusion-based merger activity. Accordingly, a naïve approach that does not correct for self-selection bias yields an insignificant ATE, and would lead to the conclusion that the type of merger being proposed does not significantly affect the rival-firm's performance outcome.

We turn now to a simple two-stage least squares estimator where we adopt a linear probability model in order to estimate the drivers of an efficiency-based merger.¹⁶ When comparing the results in the main equation for this IV procedure with the results from the naïve approach, we see that the overall results do not change substantially as Rival ROA and Acquirer Cash to Assets remain the only significant explanatory constructs. We do observe, however, that these significant explanatory variables involve a downward bias in the coefficient estimates when the naïve-OLS results are compared with the 2SLS IV results; i.e., the coefficient estimate for the Acquirer Cash to Asset is -0.013 in the naïve-OLS estimates, but is -0.018 in the IV-OLS estimates; while the Rival ROA is 0.053 in the naïve-OLS estimates, but is 0.056 in the IV-OLS estimates. Yet most importantly, the merger type (efficiency-based) variable remains insignificant, hence the ATE is still not significantly

¹⁶ We use the 'ivreg2' command in Stata, which is extremely transparent – and thus highly recommended – as it reports many diagnostics regarding the quality of instruments.

different from zero. However, the coefficient estimate for the treatment does now conform to *a priori* expectations by expressing a negative sign: i.e., efficiency-based mergers tend to harm rival firms.

It is worth underscoring that the first-stage regression for the IV approach appears to perform quite poorly, as none of the instrumental variables indicate significance. Moreover, the value statistic for a F-test for joint significance of the excluded instruments with (5,111) degrees of freedom is 0.56 with a p-value of 0.8299; thus, the joint test is not significantly different from zero which indicates weak instruments. This assessment is confirmed by additional tests of weak instrumentation: e.g., the Stock and Yogo weak ID test, a Cragg-Donald Wald F statistic of 2.21, and a Kleibergen-Paap Wald rk F statistic of 0.56. It is well known that IV estimations in the presence of weak instruments can yield coefficient estimates that involve a greater bias than simple OLS estimations (Stock and Yogo, 2005). The difficulties we experience here in terms of properly estimating the selection equation may be due to the fact that a linear regression wrongly approximates our probability model. It appears that the weaknesses of the linear probability model are quite severe in our empirical application; thus, the IV approach does not seem to be a viable alternative in order to correct for the self-selection based endogeneity present in our treatment variable.

We turn now to the maximum likelihood estimators which should be more appropriate in this empirical context, as they correctly treat the self-selection process within a non-linear probability (probit) model. Recall that the LIML estimator is characterized by being more robust as it delivers consistent – though not efficient – coefficient estimates even if the null hypothesis of joint normality on the distribution of the error terms for the substantive equation and the selection equations does not hold. But if this untestable hypothesis holds, then the FIML estimator is both consistent and more efficient as compared to the LIML estimator. Thus, the FIML estimator does potentially allow for more precise causal inferences.

When considering the empirical results based on both the LIML and FIML procedures for the various explanatory variables in the main equation of interest, the striking observation is that the results employing an FIML estimator are quite different as compared to the estimations which employ the LIML – as well as OLS – estimators. Specifically, the LIML estimator delivers results which are quite consistent with those observed when employing both the naïve-OLS and IV-OLS estimation

approaches, as only small differences regarding the size of the coefficient estimates can be observed across these three approaches. But when employing the more-efficient FIML estimator, several coefficient estimates which were previously insignificant become significant. In particular, Rival Diversification, Acquirer ROA, Acquirer Tobin's q , and Acquirer Cash to Assets—all yield significant effects once the FIML estimator is employed. In addition, the coefficient estimate for Rival ROA is now insignificantly different from zero, while it was significant in the LIML naïve-OLS and IV-OLS estimations.

Yet the most striking difference involved with employing FIML to estimate this endogenous-treatment model is that selecting an efficiency-based merger now yields in line with expectations a substantial negative effect on rival-firm value—a negative effect that is both large and significantly different from zero.¹⁷ Hence, the average treatment effect is now substantial and significantly different from zero, as the CARs of rival firms are 0.23% lower when proposed merger activity can be best characterized as efficiency-based. Moreover, this result can be interpreted as causal due to our correcting for the selection-based endogeneity involved with the treatment. It is also worth highlighting that several instrumental variables are now significant in the first stage regression when employing the FIML approach. This result contrasts with the lack of significant instruments in the 2SLS framework and the single significant instrument when employing the LIML estimator. In particular, merger characteristics (the cross-border nature of the transaction and the relative size of the target with respect to the acquirer) as well as the industry characteristics which capture the extent of the merger wave appear to significantly impact the likelihood that the merger is efficiency-based in nature.

In addition to the insignificant average treatment effect reported with the LIML estimation, an insignificant estimate for Heckman's λ (0.0293) is also reported in the LIML estimation. Heckman's λ is the product of the correlation between the error terms in both the main and selection equations ($\rho = 0.214$) and the variance of the error term in the selection equation ($\sigma =$

¹⁷ We also ran a likelihood-ratio test of the null hypothesis of no correlation between the error terms in the selection and outcome equations. Specifically, the $\chi^2(1)$ statistic takes on the value of 33.80 and the corresponding p-value is 0.0000. Therefore, we can reject the null hypothesis.

0.137). This insignificant lambda suggests that the treatment may be exogenous, as the positive correlation in Heckman's lambda is not significantly different from zero. Yet for the FIML estimator, we report the hyperbolic tangent of ρ ($\tanh \rho = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right)$) – which is estimated instead of ρ for numerical stability during optimization – and the log of σ . The corresponding values for the estimated correlation among the error terms in the substantive and selection equations (ρ) – i.e., the measure of the endogenous selection – is 0.892 and significantly different from zero. This indicates that the unobservable characteristics which affect rival-firm CARs tend to be positively correlated with the unobservable characteristics affecting the likelihood of a merger being efficiency-based in nature; i.e., endogenous selection appears to be present.

The above empirical results lead to a few observations that are generalizable to researchers doing empirical work in the context of endogenous treatment. First, taking an IV approach to dealing with issues of endogenous selection involves a few difficulties which limit the fungibility of employing this approach in a number of empirical contexts. Second, the choice between the LIML and FIML estimators is not straightforward due to the tradeoffs involved; nevertheless, FIML does involve a more-efficient estimation process. Third, and most evidently, accounting for endogenous treatment (i.e., the selection process) leads to findings that fundamentally differ from those one might obtain via a naïve approach (where the selection into treatment is assumed to be exogenous).

Endogenous Switching Regressions

We turn to the richer switching-regressions approach that allows treatment effects to go beyond an intercept effect. While the selection equation is the same as that discussed for the endogenous-treatment model – i.e., the process describing z_{jt} is represented by equation (10) – we now have two separate main equations (one for each subsample), as each outcome CAR_{ijt} is observed in only one subsample. Hence together with equation (10), we estimate the following model:

$$CAR_{ijt}^0 = \alpha_0^0 + \alpha_1^0 R_{it-1}^0 + \alpha_1^0 A_{jt-1}^0 + \varepsilon_{ijt}^0 \quad \text{if } z_{jt} = 0 \quad (11a)$$

$$CAR_{ijt}^1 = \alpha_0^1 + \alpha_1^1 R_{it-1}^1 + \alpha_2^1 A_{jt-1}^1 + \varepsilon_{ijt}^1 \quad \text{if } z_{jt} = 1 \quad (11b)$$

Unlike equation (9), we can now estimate specific effects that vary across merger type for all of the exogenous explanatory variables; i.e., we hold that $\alpha^0 \neq \alpha^1$. The variance-covariance matrix for this richer model is more complex and represented by:¹⁸

$$\Omega^s = \begin{pmatrix} \sigma_{\varepsilon^0}^2 & \rho^{01} \sigma_{\varepsilon^0} \sigma_{\varepsilon^1} & \rho^0 \sigma_{\varepsilon^0} \sigma_{\omega} \\ & \sigma_{\varepsilon^1}^2 & \rho^1 \sigma_{\varepsilon^1} \sigma_{\omega} \\ & & \sigma_{\omega}^2 \end{pmatrix}.$$

Following our discussion above, we employ two approaches to estimate this model and deal with selection bias. First, we use an endogenous switching-regressions model and simultaneously estimate (10), (11a), and (11b) together with the complete variance-covariance matrix via FIML. Second, we will partition the estimation of this model into two separate sample-selection models and apply the basic Heckman procedure. We take the latter approach for the sake of completeness and to also indicate what may be feasible in a multivariate context. The sample-selection-like approach first involves estimating the model for collusion-based mergers, i.e., equations (10) and (11a); thus, we act as if the sample was censored and we do not observe observations for the subsample where $z_{jt} = 1$. By doing so, we estimate all relevant parameters α^0 including ρ^0 and $\sigma_{\varepsilon^0}^2$. Second, we estimate the model for efficiency-based mergers, i.e., equations (10) and (11b), in a similar fashion and thus identify all of the relevant parameters α^1 including ρ^1 and $\sigma_{\varepsilon^1}^2$. As noted in our discussion of alternative techniques, this simplified approach essentially denies the existence of a potential correlation between the two error terms ε_{ijt}^0 and ε_{ijt}^1 ; i.e., it assumes $\rho^{01} = 0$. It is this strict assumption regarding correlation in the error terms along with the inefficiencies involved with such a ‘split’ regression procedure (as well as the non-ordered nature of the specified selection process) which limit the applicability of this sample-selection-like approach to self-selection issues when organizational decisions are bivariate and multivariate in nature.

¹⁸ Again we assume $\sigma_{\omega}^2 = 1$. Moreover ρ^{01} is not identified.

Endogenous Switching Results

Table 2 reports the empirical results for the switching-regressions models where selection into merger type affects how other covariates impact rival-firm CARs. In order to create a good benchmark for comparison purposes, we first describe the empirical results of a naïve approach where we simply split the sample between efficiency-based and collusion-based mergers and run separate OLS regressions. The first two columns of table 2 report the respective coefficient estimates for the efficiency-based and collusion-based subsamples. This simple split of the sample provides interesting results, as despite the smaller sample (and attendant loss in degrees of freedom) we observe some variables in the efficiency-based sample significantly impacting rival CARs. Specifically, a few acquirer characteristics (Acquirer Tobin's q , Acquirer Cash to Assets and Acquirer Diversification) affect rival-firm CARs in the context of efficiency-based mergers. The above said, we must establish whether these results are truly causal or driven by endogenous selection bias—a question we now turn to by considering the results from the approaches that take selection-bias into account.

As discussed above, different approaches exist in order to estimate a model with endogenous selection. One very feasible approach is to consider the entire problem as two separate sample-selection models and apply a standard Heckman selection approach to each of the two sub-models. Thus when analyzing efficiency-enhancing mergers, we act as if we do not observe the presence of collusion-based mergers. With such an approach, it is of course necessary to correct for the sample selection into efficiency-based mergers via a Heckman procedure. The same holds true when we focus on collusion-based merger activity. For both efficiency-based and collusion-based mergers, we estimate these models with a two-step LIML estimator as well as with a FIML estimator. As already noted, the FIML estimator hinges on the assumption of joint normality of the error terms for the substantive and selection equations. If this assumption is valid, then the FIML estimator is consistent and efficient, and LIML is consistent but not efficient. If the joint-normality assumption does not hold, then only the LIML estimator is consistent.

Akin to our empirical demonstration of endogenous treatment, the empirical results employing an LIML estimator are quite similar in terms of the size, sign and significance of coefficient estimates in the main equation as compared to the results employing a naïve-OLS approach where selection

effects are not modelled or controlled for. The similarity between the naïve-OLS sub-sampling results and the LIML results where the efficiency-based and collusion-based mergers are analyzed via a sample-selection-like approach holds for both the efficiency-based and collusion-based merger types. The above said, some coefficient estimates with the naïve approach do appear to suffer a minimal degree of bias. We should also note that only one excluded instrument (cross-border mergers) is significant in the two LIML selection equations, though a number of included instruments – i.e., regressors from the substantive equation – indicate statistical significance in the selection equation. Most importantly, the coefficient estimate for the inverse Mill's ratio (Heckman's lambda) is not significant in the LIML estimations.

Also similar to our demonstration of endogenous-treatment analysis, the FIML estimations that take a sample-selection-like approach deliver results that substantially differ from the naïve-OLS and LIML sample-selection-like approaches. In particular, the coefficient estimates for a number of regressors (Rival ROA, Rival Market Capitalization, Rival Diversification, Acquirer ROA, Acquirer Tobin's q, and Acquirer Cash to Assets) in the main equation for efficiency-based mergers are now significantly different from zero. But even more importantly, we now estimate a positive correlation (significant at the 1% level) between the error terms of the substantive and selection equations (ρ^1)—a result which strongly suggests the presence of endogeneous selection.¹⁹

In order to measure the treatment effect in this model where we interact the treatment variable with all of the other regressors, we would need to calculate its marginal effect at, say, the mean (or another) value for all of the regressors. Doing this separately for each subsample, however, would produce the so-called average treatment effect on the treated (ATET) rather than the ATE.²⁰ Given the

¹⁹ For numerical stability in the optimization problem, the FIML estimator estimates the inverse hyperbolic tangent of ρ . The corresponding estimated value for the correlation is 0.976. Also in this case we run a likelihood ratio test. We can strongly reject the null hypotheses of lack of correlation between the error terms in the selection equations and the substantive equations in both subsamples. Specifically, the $\chi^2(1)$ statistics take on the value of 14.55 and 24.70 –with corresponding p-values of 0.0001 and 0.0000— for the efficiency-based mergers and the collusion-based mergers respectively.

²⁰ When we interact the treatment variable with other covariates we can estimate the ATET, which measures the average difference between the outcomes for the treatment and the control group *based on the treated population*; i.e., when the treatment dummy is equal to 1 (or to 0). The ATET may differ from the ATE because

large number of regressors, we do not report this effect. But we can compare the sign and size of the selection correction obtained via this approach with that obtained via the simpler endogenous treatment model discussed above. Heckman's lambda – i.e., the coefficient measuring the impact of endogenous selection – is obtained by taking the product between the correlation among the errors terms in the selection and substantive equations (ρ^i , $i = 0,1$) and the standard deviation of the latter (σ_{ε^i} , $i = 0,1$). In the efficiency-based sample, for instance, this term would be negative. This result should be interpreted as indicating that rival firms facing an efficiency-based merger earn significantly lower CARs as compared to a potentially random firm from the same sample would have earned.

Moving to the results for the subsample of collusion-based merger activity, we obtain qualitatively similar results when we focus on the subsample of collusion-based mergers while employing the sample-selection-like approach to dealing with selection into these mergers. While the empirical results in the main equation where we use an FIML estimator differ from the coefficient estimates yielded by the naïve sub-sampling approach and the LIML estimator approach, it is fair to point out that these differences regarding collusion-based mergers are not as substantial as compared to efficiency-based mergers. We also estimate a significant correlation between the error terms in the substantive and selection equations in this case; moreover, this correlation is negative as expected, since the correlation in the efficiency-based subsample is positive. Similar to the interpretation noted above, this finding indicates that rival firms facing a collusion-based merger earn significantly higher CARs as compared to what a potentially random firm from the same sample might have earned.

Finally, we present the FIML estimates for an endogenous switching regression model in the last three columns of Table 2. The expectation for this type of model is that it should deliver more-efficient estimates as compared to the approach where two distinct Heckman FIML models are employed to consider efficiency-based and collusion-based mergers separately. The switching-regressions model simultaneously employs all of the available information and imposes the restriction that the coefficient for the selection into the two equations (efficiency-based and collusion-based) is

the effect of the treatment varies over the values of the covariates with which it is interacted. ATET and ATE tend to be similar if the average predicted outcome for the treated group is similar to that for the whole population.

the same. Additionally, this estimator makes an even stronger assumption on the distribution of the error terms (the error terms of the two main equations and the error term in the selection equation), as it is assumed to be tri-variate normally distributed.

Our empirical results for the switching-regressions model are in line with the results where we partition the estimation into two separate sample-selection models and estimate via FIML. First, consistent with our expectations, the coefficient estimates appear to be estimated a bit more precisely under the switching-regressions approach as the standard errors are generally a bit smaller than those obtained when employing two-separate Heckman procedures. Furthermore, we still observe that the treatment involves significant slope effects in the sense that the coefficient estimates for efficiency-based mergers α^1 are different from the coefficient estimates for collusion-based mergers α^0 . Second, the excluded instruments in the selection equation seem to work much better in this model as compared to the approach where we employ two separate Heckman procedures. This result is due to the greater efficiency involved with there now being only one selection equation, whereas the sample-selection-like FIML approach involved two separate equations. Third, the estimates for the ancillary parameters (ρ^i and σ_{ε^i} , $i = 0,1$) are very similar to those obtained above, thus supporting the finding that there is significant endogenous selection into both subsamples.²¹

The different approaches demonstrated above all allow the self-selected treatment to involve differences in the relevant coefficient estimates for the other explanatory variables in the substantive main equation. Allowing for endogenous switching certainly adds to the endogenous treatment approach which was previously demonstrated, as the impact of the treatment can go beyond a simple intercept effect in many empirical contexts. In our specific context, it does appear that acquirer and rival characteristics involve heterogeneous effects on rival-firm CARs that depend on the type of merger selected into. Furthermore, our empirical demonstration also supports the prior that Lee's (1978, 1982) endogenous switching-regressions procedure – which was also advocated by Hamilton and Nickerson (2003) – represents an efficient means to estimate a full model in a context with

²¹ Also in this case, the likelihood ratio test of independent equations rejects the null hypothesis. Specifically the p-value for the $\chi^2(2) = 36.85$ (remember that we now have two degrees of freedom since we test two restrictions) is equal to 0.0000.

endogenous-switching. Nevertheless, the less cumbersome sample-selection-like approach to this issue does yield comparable results and also represents an approach that may be applicable in an environment involving multivariate organizational decisions.

Conclusions and Prescriptions

Our study first provides a clear caution to empirical researchers throughout management that neglecting the endogenous nature of the selection into a particular organizational decision can lead to biased and inconsistent empirical results. Moreover, such biased results lead to spurious causal inferences that prevent the generation of well-informed theoretical and managerial conclusions. Taking both an informal and formal approach, we outline the exact nature of self-selection based endogeneity bias. Furthermore, our empirical demonstration illustrates how employing a naïve-OLS approach in a context where endogenous treatment and endogenous switching are present can lead to biased and erroneous findings. While the management literature has begun to appreciate over the last decade the importance of factoring self-selection effects due to the prevalence of this issue in much of our empirical scholarship, it is nevertheless important to continue to remind, outline and demonstrate to researchers that they need to be conscious of selection-based biases in their empirical estimations.

Beyond the need for researchers to correctly recognize the existence of selection-based endogeneity, it is also important that researchers clearly comprehend how this endogeneity manifests; i.e., what is the exact source of the endogeneity bias. It is, of course, the self-selection by managers into different organizational decisions which represents the principal channel via which omitted factors raise the issue of self-selection based endogeneity bias in the management discipline. Indeed in many management applications, organizational decisions tend to be discrete rather than continuous; hence, the researcher must correctly model this discrete process in order to deal with the selection effects. Basile (2008) and others have advocated the use of simple linear-probability models in order to estimate selection equations; i.e., the process which explains the manifestation of the treatment. Our review and demonstration, however, suggest that employing such an approach might lead to misleading results. Based on our own empirical application, for instance, employing a linear estimator via 2SLS would generate incorrect inferences and conclusions. In particular, the selection equation in

our empirical demonstration indicated mis-specification as the instruments did not appear to work correctly. Moreover, our review of the methodological literature suggests that relying on a linear approximation in order to model a discrete treatment variable often runs into similar problems. It is comforting then that our review of the literature suggests that this approach has not been commonly taken up by empirical researchers in management. Yet it is worth mentioning the limitations of this approach so that future researchers continue to be cautious with respect to employing a linear approximation to model the selection into a treatment.

In light of selection-bias ultimately being a sub-set of endogeneity bias, a fundamental issue of concern for empirical researchers in management is the identification of appropriate exclusion restrictions – i.e., instruments – in order to tackle the endogeneity problem. Unfortunately, this issue still seems to be a major problem in empirical research in management. In particular, many empirical studies appear to employ instruments that are not truly exogenous as they correlate with the error term in the substantive equation. Furthermore, many studies suffer from weak instrumentation in the sense that the instruments do not significantly correlate with the potentially endogenous variable. Such practices can lead to coefficient estimates that involve greater biases as compared to results employing naïve OLS estimates. Even more troubling is the common practice by researchers to be quite un-transparent regarding this issue, as the selection-equation estimations are often not reported. In addition to not spending adequate time to outline the implicit identification strategy, diagnostics which test the quality of the instruments are often not mentioned or discussed. It is also worth mentioning that the choice of instruments and the choice of methodological approach can interact and play a crucial role in empirical applications. For example, the excluded instruments in our 2SLS linear-approximation application do not indicate correlation with treatment despite their resting on theoretical foundations. But when employing the more appropriate non-linear approaches, our instruments work quite well and we generate significant selection effects. Thus, concluding from the 2SLS approach that the instruments are generally weak would be incorrect, as the real issue was the inappropriate specification of the selection equation. We strongly urge then that empirical scholars spend considerably more time, care and attention with regard to properly identifying their selection equations, as far too often the approach to this crucial issue seems to be cavalier and superficial.

The empirical researcher must also be quite careful in setting up the correct modelling approach to deal with the exact nature of the endogenous self-selection and how the treatment variable affects the outcome of ultimate interest. In particular, it is important in the context of self-selection based endogeneity to differentiate between endogenous treatment (which involves an intercept effect) and endogenous switching (which involves different slope coefficients for the other explanatory variables). Under endogenous treatment, a researcher may be convinced (or more interested) that the treatment effect manifests as an intercept effect. In this case, one assumes that all other covariates similarly affect the outcome variable independent of the treatment; hence, the treatment simply shifts the regression line up or down. Under endogenous switching, however, a researcher holds that the treatment effect is mediated by the other explanatory variables and thus also involves a slope effect. Moreover, these two variants (endogenous treatment and endogenous switching) of self-selection bias are distinct from the foundational sample-selection bias. Yet the differences between endogenous treatment, endogenous switching and sample-selection are quite subtle; thus, these differences have not been fully appreciated in the empirical literature and researchers far too often mix-up these different types of selection effects. Since the alternative methodological approaches line up with the different ways via which selection can manifest, the choice of modelling approach should be principally guided by the specific research question being faced by the researcher.

While a few different modelling approaches exist to deal with the different forms of selection-bias, the empirical researcher mindful of selection-based endogeneity can also choose between different estimation methods. Characteristic of the different selection models is that the selection process is generally modeled as a probit, the error term in the selection equation is commonly assumed to be normally distributed, and the alternative estimators are based on maximum likelihood. In particular, we extensively discussed the choice between the LIML and FIML estimators in the context of full selection models that entail a selection equation into the treatment as well as one or more main substantive equations. The LIML estimator is more robust but also less efficient as compared to the FIML estimator. FIML is asymptotically efficient, but it hinges on the additional assumption of joint normality between the error terms of the selection equation and the error term(s) of the substantive equation(s). While there is no clear-cut rule with respect to choosing between these two estimators, it

is fair to state that the additional assumption residing behind the FIML estimator becomes less restrictive when the sample gets larger. Thus, the LIML estimator is potentially advisable when the sample is relatively small. Moreover, our main suggestion regarding this issue is that researchers be cautious and transparent regarding the choice between LIML and FIML; i.e., discuss the tradeoffs involved with the two estimators, and show the empirical results while employing both estimators.

Selection into organizational decisions is not a randomized process but is instead an endogenous process undertaken by managers with the intent to enhance organizational outcomes. Since empirical researchers in management often study the effect of these discrete organizational choices on outcomes by employing observational data sets, the issue of unobservable factors which correlate with both the organizational decision and the outcome variable of interest indicates that self-selection based endogeneity is both fundamental and characteristic of much of what we study in management. Our contribution here attempts to clarify the nature of selection-based endogeneity bias and provide researchers in management with a transparent understanding of the different methodologies available in order to correctly model and control for selection-effects. Our hope is that this effort might spur researchers in management to be more comprehensive and precise in dealing with selection-effects in their empirical applications. Enhanced efforts in our discipline to think carefully and be transparent about the sources of endogeneity in our empirical models will lead to sounder causal inferences that are both consistent and unbiased, which in turn ultimately leads to stronger theoretical and managerial conclusions: the ultimate test of our discipline.

References

- Andrade, G., Mitchell, M., & Stafford, E. (2001). New evidence and perspectives on mergers. *Journal of Economic Perspectives*, 15, 103–120.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120.
- Bascle, G. (2008). Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organization*, 6(3), 285–327.
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48(3), 386–398.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1), 1–51.
- Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2014). Does working from home work? Evidence from a Chinese experiment. Working paper, Stanford University.
- Brouthers, K. D., Brouthers, L. E., & Werner, S. (2003). Transaction cost-enhanced entry mode choices and firm performance. *Strategic Management Journal*, 24(12), 1239–1248.
- Capron, L., & Shen, J. C. (2007). Acquisitions of private vs. public firms: Private information, target selection, and acquirer returns. *Strategic Management Journal*, 28(9), 891–911.
- Chatterjee, S. (1992). Sources of value in takeovers: Synergy or restructuring-implications for target and bidder firms. *Strategic Management Journal*, 13(4), 267–286.
- Clougherty, J. A., & Duso, T. (2011). Using rival effects to identify synergies and improve merger typologies. *Strategic Organization*, 9(4), 310–335.
- Duso, T., Gugler, K., & Szücs, F. (2013). An empirical assessment of the 2004 EU merger policy reform. *Economic Journal*, 123(572), F596–F619.
- Duso, T., Gugler, K., & Yurtoglu, B. (2010). Is the event study methodology useful for merger analysis: A comparison of stock market and accounting data. *International Review of Law and Economics*, 30(2), 186–192.
- Duso, T., Neven, D. J., & Röller, L. H. (2007). The political economy of European merger control: Evidence using stock market data. *Journal of Law and Economics*, 50(3), 455–489.
- Echambadi, R., Campbell, B., & Agarwal, R. (2006). Encouraging best practice in quantitative management research: An incomplete list of opportunities. *Journal of Management Studies*, 43(8), 1801–1820.
- Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, 1(1), 51–78.
- Harford, J. (2005). What drives merger waves? *Journal of Financial Economics*, 77(3), 529–560.
- Heckman, J. J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42(4), 679–694.
- Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5(4), 475–492.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.
- Hitt, M. A., Harrison, J. S., & Ireland, R. D. (2001). *Mergers and acquisitions: A guide to creating value for shareholders*. Oxford: Oxford University Press.
- King, A. A., & Tucci, C. L. (2002). Incumbent entry into new market niches: The role of experience and managerial choice in the creation of dynamic capabilities. *Management Science*, 48(2), 171–186.
- Lee, L. F. (1978). Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *International Economic Review*, 19(2), 415–433.
- Lee, L. F. (1979). Identification and estimation in binary choice models with limited (censored) dependent variables. *Econometrica*, 47(4), 977–996.
- Lee, L. F. (1982). Some approaches to the correction of selectivity bias. *The Review of Economic Studies*, 49(3), 355–372.
- Li, M. (2012). Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods*, 16(2), 188–226.
- Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

- Mahoney, J. M., & Mahoney, J. T. (1993). An empirical investigation of the effect of corporate charter antitakeover amendments on stockholder wealth. *Strategic Management Journal*, 14(1), 17-31.
- McWilliams, A., & Siegel, D. (1997). Event study in management research: Theoretical and empirical issues. *Academy of Management Journal*, 40(3), 626-657.
- Modrek, S., & Cullen, M. R. (2013). Health consequences of the 'Great Recession' on the employed: Evidence from an industrial cohort in aluminum manufacturing. *Social Science & Medicine*, 92, 105-113.
- Morrow, J. L., Sirmon, D. G., Hitt, M. A., & Holcomb, T. R. (2007). Creating value in the face of declining performance: Firm strategies and organizational recovery. *Strategic Management Journal*, 28(3), 271-283.
- Robins, J., & Wiersema, W. 1995. A resource-based approach to the multi-business firm: Empirical analysis of portfolio inter-relationships and corporate financial performance. *Strategic Management Journal*, 16, 277-300.
- Scholes, M., & Williams, J. (1977). Estimating betas from non-synchronous data. *Journal of Financial Economics*, 5(3), 309-327.
- Seth, A. (1990). Value creation in acquisitions: A re-examination of performance issues. *Strategic Management Journal*, 11(2), 99-115.
- Shaver, J. M. (1998). Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival. *Management Science*, 44(4), 571-585.
- Stock, J., & Yogo, M. (2005) Testing for weak instruments in linear IV regression. In: Andrews D.W.K. (Eds.), *Identification and Inference for Econometric Models* (pp. 80-108). New York: Cambridge University Press.
- Turk, T. A. (1992). Takeover resistance, information leakage, and target firm value. *Journal of Management*, 18(3), 503-522.
- Villalonga, B., & McGahan, A. (2005). The choice among acquisitions, alliances, and divestitures. *Strategic Management Journal*, 26(13), 1183-1208.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.
- Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach*. Mason, OH: Thomson-South Western.

Table 1. Endogenous Treatment Estimations.

| | Exogenous | Endogenous - linear | | Endogenous - Non-linear | | Endogenous - Non-linear | |
|--------------------------------|--------------------|---------------------|----------------------|-------------------------|-----------------------|-------------------------|-----------------------|
| | OLS | IV | | LIML | | FIML | |
| | Main | Main | First Stage | Main | Selection | Main | Selection |
| | CAR(50,5) | CAR(50,5) | Efficiency=1 | CAR(50,5) | Efficiency=1 | CAR(50,5) | Efficiency=1 |
| Rival ROA | 0.053* (0.030) | 0.056* (0.031) | -0.010 (0.120) | 0.0549* (0.0314) | -0.102 (0.320) | 0.0605 (0.0397) | -0.0782 (0.322) |
| Rival Tobin's Q | 0.003 (0.013) | 0.003 (0.013) | -0.006 (0.030) | 0.00281 (0.00884) | -0.0133 (0.0942) | 0.00178 (0.0112) | -0.0347 (0.0825) |
| Rival Cash to Assets | -0.004 (0.004) | -0.004 (0.004) | -0.005 (0.012) | -0.00379 (0.00401) | -0.0177 (0.0429) | -0.00383 (0.00508) | -0.0218 (0.0395) |
| Rival Debt to Total Assets | 0.020 (0.069) | 0.033 (0.071) | 0.052 (0.169) | 0.0282 (0.0577) | 0.258 (0.590) | 0.0654 (0.0718) | 0.886 (0.543) |
| Rival Market Capital. | -0.006 (0.006) | -0.007 (0.006) | -0.011 (0.018) | -0.00672 (0.00513) | -0.0210 (0.0518) | -0.00961 (0.00636) | 0.0546 (0.0483) |
| Rival Diversification | 0.002 (0.003) | 0.004 (0.005) | 0.034*** (0.013) | 0.00347 (0.00396) | 0.111*** (0.0362) | 0.0101** (0.00434) | 0.0941*** (0.0320) |
| Rival Relative Size to Acq. | 0.002 (0.003) | 0.002 (0.003) | -0.005 (0.013) | 0.00243 (0.00256) | -0.0186 (0.0276) | 0.00321 (0.00322) | -0.0287 (0.0252) |
| Acq. ROA | -0.019 (0.031) | -0.048 (0.044) | -0.370* (0.191) | -0.0367 (0.0414) | -1.155*** (0.401) | -0.114*** (0.0433) | -0.881** (0.346) |
| Acq. Tobin's Q | 0.004 (0.006) | 0.007 (0.009) | 0.066 (0.059) | 0.00553 (0.00665) | 0.186*** (0.0580) | 0.0137* (0.00771) | 0.172*** (0.0521) |
| Acq. Cash to Asset | -0.013* (0.007) | -0.018* (0.010) | -0.072 (0.043) | -0.0160* (0.00824) | -0.216*** (0.0708) | -0.0314*** (0.00870) | -0.221*** (0.0638) |
| Acq. Debt to total Asset | 0.071 (0.063) | 0.061 (0.067) | 0.086 (0.392) | 0.0654 (0.0636) | 0.410 (0.691) | 0.0468 (0.0800) | -0.119 (0.613) |
| Acq. Market Capital. | 0.009 (0.008) | 0.008 (0.008) | -0.009 (0.056) | 0.00846 (0.00722) | -0.00931 (0.0571) | 0.00844 (0.00814) | -0.0311 (0.0514) |
| Acq. Diversification | 0.005 (0.004) | 0.004 (0.004) | -0.014 (0.028) | 0.00431 (0.00377) | -0.0505 (0.0373) | 0.000663 (0.00458) | -0.0470 (0.0346) |
| Efficiency-based merger=1 | 0.017 (0.014) | -0.059 (0.075) | | -0.0303 (0.0638) | | -0.233*** (0.0214) | |
| Merger Relatedness | | | 0.0411 (0.124) | | 0.177 (0.217) | | 0.0262 (0.141) |
| Merger Cross-border | | | 0.182 (0.112) | | 0.591*** (0.158) | | 0.283*** (0.108) |
| Merger Relative Size Tar./Acq. | | | 0.00396 (0.0130) | | 0.0156 (0.0233) | | 0.0337** (0.0167) |
| Notifications | | | -0.259 (1.108) | | -0.364 (1.505) | | -2.232** (1.023) |
| Notifications squared | | | 0.107 (0.836) | | -0.0640 (1.187) | | 1.437* (0.824) |
| Sales growth | | | 0.114 (0.249) | | 0.345 (0.535) | | 0.200 (0.382) |
| Share return | | | 0.0138 (0.0647) | | 0.0264 (0.0955) | | -0.0368 (0.0629) |
| Tobin's q | | | 0.000889 (0.0196) | | -0.00128 (0.0285) | | 0.0261 (0.0207) |
| Liberalization | | | 0.0515 (0.0592) | | 0.141 (0.0917) | | 0.0970 (0.0602) |
| Lambda | | | | 0.0293 (0.0386) | | | |
| ρ (ln) | | | | | | 1.430*** (0.140) | |
| σ (ln) | | | | | | -1.748*** (0.0479) | |
| Constant | -0.0841 (0.101) | -0.00718 (0.111) | 0.696 (0.836) | -0.0422 (0.106) | | 0.0988 (0.0892) | |
| N | 431 | 431 | 431 | 431 | 431 | 431 | 431 |

Table 2. Endogenous Switching Estimations.

| | Exogenous | | Sample-Selection-Like | | Sample-Selection-Like | | Sample-Selection-Like | | Sample-Selection-Like | | Endogenous Switching Regressions | | |
|-----------------------------------|-------------------------|-----------------------|------------------------|-----------------------|------------------------|------------------------|------------------------|-----------------------|-----------------------|-----------------------|----------------------------------|------------------------|----------------------|
| | OLS | | LIML | | FIML | | LIML | | FIML | | FIML | | |
| | Efficiency | Collusion | Efficiency | | Efficiency | | Collusion | | Collusion | | Efficiency | Collusion | Selection |
| | Main | Main | Main | Selection | Main | Selection | Main | Selection | Main | Selection | Main | Main | Selection |
| | CAR(50,5) | CAR(50,5) | CAR(50,5) | Efficiency=1 | CAR(50,5) | Efficiency=1 | CAR(50,5) | Collusive=1 | CAR(50,5) | Collusive=1 | CAR(50,5) | CAR(50,5) | Efficiency |
| Rival ROA | 0.0594 (0.0611) | 0.0432 (0.0399) | 0.0686 (0.0598) | -0.102 (0.320) | 0.119* (0.0673) | -0.123 (0.306) | 0.0425 (0.0403) | 0.102 (0.320) | 0.0250 (0.0461) | -0.214 (0.367) | 0.149** (0.0715) | 0.0219 (0.0464) | 0.190 (0.365) |
| Rival Tobin's Q | -0.0125 (0.0153) | 0.00867 (0.0110) | -0.0128 (0.0147) | -0.0133 (0.0942) | -0.0162 (0.0182) | -0.0320 (0.0939) | 0.00787 (0.0111) | 0.0133 (0.0942) | 0.0110 (0.0125) | 0.0782 (0.0868) | -0.0283* (0.0172) | 0.0120 (0.0126) | -0.095 (0.090) |
| Rival Cash to Assets | 0.00712 (0.00703) | -0.00631 (0.00501) | 0.00717 (0.00673) | -0.0177 (0.0429) | 0.00713 (0.00814) | -0.0265 (0.0401) | -0.00614 (0.00506) | 0.0177 (0.0429) | -0.00821 (0.00572) | -0.0167 (0.0423) | 0.0128 (0.00819) | -0.00952* (0.00573) | -0.0003 (0.041) |
| Rival Debt to total Assets | 0.0220 (0.100) | 0.00610 (0.0703) | 0.0316 (0.0969) | 0.258 (0.590) | 0.0223 (0.118) | 0.852 (0.564) | 0.0210 (0.0721) | -0.258 (0.590) | 0.0406 (0.0798) | -0.663 (0.569) | 0.0513 (0.113) | 0.0734 (0.0803) | 1.056* (0.565) |
| Rival Market Capital. | -0.0100 (0.00880) | -0.00582 (0.00627) | -0.0122 (0.00902) | -0.0210 (0.0518) | -0.0218* (0.0121) | 0.0116 (0.0560) | -0.00686 (0.00639) | 0.0210 (0.0518) | -0.0105 (0.00716) | -0.0241 (0.0498) | -0.0146 (0.0108) | -0.00969 (0.00721) | 0.051 (0.053) |
| Rival Diversification | 0.00698 (0.00582) | 0.00247 (0.00429) | 0.00935 (0.00657) | 0.111*** (0.0362) | 0.0183*** (0.00696) | 0.0954*** (0.0338) | 0.00568 (0.00513) | -0.111*** (0.0362) | 0.0109** (0.00491) | -0.103*** (0.0329) | 0.0157** (0.00669) | 0.0105** (0.00486) | 0.094*** (0.032) |
| Rival Relative Size to Acq. | -0.00002 (0.00439) | 0.00430 (0.00319) | 0.000323 (0.00424) | -0.0186 (0.0276) | 0.00474 (0.00606) | -0.0280 (0.0277) | 0.00445 (0.00321) | 0.0186 (0.0276) | 0.00382 (0.00361) | 0.0172 (0.0253) | 0.00425 (0.00576) | 0.00301 (0.00371) | -0.0324 (0.027) |
| Acq. ROA | -0.0516 (0.0598) | -0.00454 (0.0485) | -0.0847 (0.0750) | -1.155*** (0.401) | -0.231*** (0.0791) | -0.973*** (0.375) | -0.0321 (0.0545) | 1.155*** (0.401) | -0.0730 (0.0541) | 1.024*** (0.362) | -0.215*** (0.0775) | -0.0732 (0.0536) | -0.932*** (0.352) |
| Acq. Tobin's Q | 0.0198* (0.0105) | -0.00697 (0.00824) | 0.0216** (0.0103) | 0.186*** (0.0580) | 0.0254** (0.0115) | 0.204*** (0.0545) | -0.000640 (0.00985) | -0.186*** (0.0580) | 0.0127 (0.00906) | -0.160*** (0.0521) | 0.0268** (0.0113) | 0.0150 (0.00929) | 0.195*** (0.055) |
| Acq. Cash to Asset | -0.0305*** (0.00961) | -0.00258 (0.0106) | -0.0357*** (0.0120) | -0.216*** (0.0708) | -0.0573*** (0.0146) | -0.212*** (0.0702) | -0.00790 (0.0115) | 0.216*** (0.0708) | -0.0124 (0.0112) | 0.231*** (0.0668) | -0.0535*** (0.0137) | -0.0162 (0.0110) | -0.192*** (0.075) |
| Acq. Debt to total Asset | 0.0662 (0.107) | 0.112 (0.0804) | 0.0574 (0.104) | 0.410 (0.691) | 0.118 (0.143) | -0.227 (0.723) | 0.101 (0.0817) | -0.410 (0.691) | 0.0888 (0.0921) | 0.0332 (0.648) | 0.0531 (0.127) | 0.0541 (0.0942) | -0.542 (0.716) |
| Acq. Market Capital. | 0.0154 (0.0119) | 0.0127 (0.00963) | 0.0166 (0.0114) | -0.00931 (0.0571) | 0.0282* (0.0158) | -0.0000989 (0.0603) | 0.00859 (0.0102) | 0.00931 (0.0571) | -0.00212 (0.0105) | 0.0352 (0.0537) | 0.0195 (0.0143) | -0.00222 (0.0110) | -0.086 (0.083) |
| Acq. Diversification | 0.0148** (0.00589) | -0.00113 (0.00477) | 0.0139** (0.00582) | -0.0505 (0.0373) | 0.00792 (0.00698) | -0.0271 (0.0351) | -0.00237 (0.00491) | 0.0505 (0.0373) | -0.00223 (0.00530) | 0.0465 (0.0349) | 0.00697 (0.00668) | -0.000704 (0.00528) | -0.024 (0.035) |
| Relatedness | | | | 0.177 (0.217) | | -0.125 (0.138) | | -0.177 (0.217) | | -0.239 (0.157) | | | -0.007 (0.149) |
| Merger Cross-border | | | | 0.591*** (0.158) | | 0.387*** (0.110) | | -0.591*** (0.158) | | -0.409*** (0.125) | | | 0.297*** (0.105) |
| Merger Relative Size Tar./Acq. | | | | 0.0156 (0.0233) | | 0.0239 (0.0169) | | -0.0156 (0.0233) | | -0.0291* (0.0171) | | | 0.028* (0.016) |
| Notifications | | | | -0.364 | | -1.841 | | 0.364 | | 0.924 | | | -1.955* |

| | | | | | | | | | | | | | |
|-----------------------|---------|----------|---------|----------|-----------|----------|---------|----------|-----------|-----------|-----------|-----------|---------|
| | | | | (1.505) | | (1.204) | | (1.505) | | (1.179) | | (1.038) | |
| Notifications squared | | | | -0.0640 | | 1.220 | | 0.0640 | | -0.309 | | 1.306 | |
| | | | | (1.187) | | (0.965) | | (1.187) | | (0.942) | | (0.851) | |
| Sales growth | | | | 0.345 | | -0.256 | | -0.345 | | -0.385 | | -0.041 | |
| | | | | (0.535) | | (0.537) | | (0.535) | | (0.424) | | (0.472) | |
| Share return | | | | 0.0264 | | -0.0573 | | -0.0264 | | 0.00523 | | -0.030 | |
| | | | | (0.0955) | | (0.0901) | | (0.0955) | | (0.0660) | | (0.078) | |
| Tobin's q | | | | -0.00128 | | 0.00302 | | 0.00128 | | -0.0209 | | 0.027 | |
| | | | | (0.0285) | | (0.0203) | | (0.0285) | | (0.0210) | | (0.020) | |
| Liberalization | | | | 0.141 | | 0.00151 | | -0.141 | | -0.169** | | 0.053 | |
| | | | | (0.0917) | | (0.0628) | | (0.0917) | | (0.0687) | | (0.086) | |
| Lambda | | 0.0364 | | | | | | -0.0704 | | | | | |
| | | (0.0535) | | | | | | (0.0604) | | | | | |
| σ^1 (ln) | | | | | -1.637*** | | | | | | -1.687*** | | |
| | | | | | (0.149) | | | | | | (0.173) | | |
| σ^0 (ln) | | | | | | | | | -1.732*** | | -1.732*** | | |
| | | | | | | | | | (0.0580) | | (0.0597) | | |
| ρ^1 (ln) | | | | | 2.206*** | | | | | | 1.934** | | |
| | | | | | (0.752) | | | | | | (0.778) | | |
| ρ^0 0 (ln) | | | | | | | | | | -1.672*** | | -1.597*** | |
| | | | | | | | | | (0.266) | | (0.271) | | |
| Constant | -0.161 | -0.143 | -0.175 | | -0.333** | | -0.0184 | | 0.245* | | -0.274 | 0.239* | 0.710 |
| | (0.155) | (0.121) | (0.148) | | (0.160) | | (0.159) | | (0.130) | | (0.174) | (0.142) | (1.030) |
| N | 136 | 295 | 431 | 136 | 431 | 136 | 431 | 295 | 431 | 295 | 431 | 431 | 431 |
| Censored | | | 136 | | 136 | | 295 | | 295 | | | | |

Appendix A. Sample-Selection Studies in SMJ (2002-2014).

| Article | Correction Method | Reported Selection Equation | Identification Variable(s) | Significant Selection | Self Selection Present |
|--|--|-----------------------------|----------------------------|-----------------------|---------------------------|
| Baum, J. A., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. <i>Strategic Management Journal</i> , 21(3), 267-294. | Auxiliary, Lee's generalization of Heckman | | | | |
| Chang, S. J., & Rosenzweig, P. M. (2001). The choice of entry mode in sequential foreign direct investment. <i>Strategic Management Journal</i> , 22(8), 747-776. | Omits observations | None | | | |
| Canyon, M. J., Peck, S. I., & Sadler, G. V. (2001). Corporate tournaments and executive compensation: Evidence from the UK. <i>Strategic Management Journal</i> , 22(8), 805-815. | Tests for selection | No | Yes | No | |
| Westphal, J. D., & Fredrickson, J. W. (2001). Who directs strategic change? Director experience, the selection of new CEOs, and change in corporate strategy. <i>Strategic Management Journal</i> , 22(12), 1113-1137. | Heckman 2-stage | No | Not Reported | Not Reported | Evident but not discussed |
| Gulati, R., & Higgins, M. C. (2003). Which ties matter when? The contingent effects of interorganizational partnerships on IPO success. <i>Strategic Management Journal</i> , 24(2), 127-144. | Heckman 2-stage | No | Yes | No | |
| Delmar, F., & Shane, S. (2003). Does business planning facilitate the development of new ventures? <i>Strategic Management Journal</i> , 24(12), 1165-1185. | Lee's generalization of Heckman | Yes | Yes | Yes | |
| Shamsie, J., Phelps, C., & Kuperman, J. (2004). Better late than never: A study of late entrants in household electrical equipment. <i>Strategic Management Journal</i> , 25(1), 69-84. | Heckman 2-stage | Yes | No | No | |
| Sanders, W. M., & Boivie, S. (2004). Sorting things out: Valuation of new firms in uncertain markets. <i>Strategic Management Journal</i> , 25(2), 167-186. | Heckman 2-stage | No | No | No | |
| Huyghebaert, N., & Van de Gucht, L. M. (2004). Incumbent strategic behavior in financial markets and the exit of entrepreneurial start-ups. <i>Strategic Management Journal</i> , 25(7), 669-688. | Heckman 2-stage | No | Not Clear | No | |

| | | | | | |
|--|---------------------------------|---------------|-----------|--------------|---------------------------|
| Mishina, Y., Pollock, T. G., & Porac, J. F. (2004). Are more resources always better for growth? Resource stickiness in market and product expansion. <i>Strategic Management Journal</i> , 25(12), 1179-1197. | Heckman 2-stage | No | No | No | Evident but not discussed |
| Haas, M. R., & Hansen, M. T. (2005). When using knowledge can hurt performance: The value of organizational capabilities in a management consulting company. <i>Strategic Management Journal</i> , 26(1), 1-24. | Alternative 2-part Logit model | Yes | Yes | No | |
| Nachum, L., & Zaheer, S. (2005). The persistence of distance? The impact of technology on MNE motivations for foreign investment. <i>Strategic Management Journal</i> , 26(8), 747-767. | Fixed industry effects | There is None | N/A | N/A | |
| Park, N. K., & Mezas, J. M. (2005). Before and after the technology sector crash: The effect of environmental munificence on stock market response to alliances of e-commerce firms. <i>Strategic Management Journal</i> , 26(11), 987-1007. | Omits observations | There is None | N/A | N/A | Evident but not discussed |
| Higgins, M. C., & Gulati, R. (2006). Stacking the deck: The effects of top management backgrounds on investor decisions. <i>Strategic Management Journal</i> , 27(1), 1-25. | Heckman 2-stage | No | No | Not Reported | Evident but not discussed |
| Westphal, J. D., Boivie, S., Chng, M., & Han, D. (2006). The strategic impetus for social network ties: Reconstituting broken CEO friendship ties. <i>Strategic Management Journal</i> , 27(5), 425-445. | Heckman 2-stage | No | Not Clear | Not Reported | |
| Henderson, A. D., Miller, D., & Hambrick, D. C. (2006). How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance. <i>Strategic Management Journal</i> , 27(5), 447-460. | Lee's generalization of Heckman | No | Not Clear | Yes | |
| Hayward, M. L., & Shimizu, K. (2006). De-commitment to losing strategic action: Evidence from the divestiture of poorly performing acquisitions. <i>Strategic Management Journal</i> , 27(6), 541-557. | Auxiliary Heckman 2-stage | No | Not Clear | Not Reported | Evident but not discussed |
| Wright, P., Kroll, M., Krug, J. A., & Pettus, M. (2007). Influences of top management team incentives on firm risk taking. <i>Strategic Management Journal</i> , 28(1), 81-89. | Auxiliary Logit Estimation | No | No | No | |
| Karaevli, A. (2007). Performance consequences of new CEO 'Outsiderness': Moderating effects of pre-and post-succession contexts. <i>Strategic Management Journal</i> , 28(7), 681-706. | Heckman 2-stage | No | Yes | No | Evident but not discussed |
| Nerkar, A., & Shane, S. (2007). Determinants of invention commercialization: An empirical examination of academically sourced inventions. <i>Strategic Management Journal</i> , 28(11), 1155-1166. | Lee's generalization of Heckman | No | Yes | No | |

| | | | | | |
|--|--|-----|-----------|---------------------------|---------------------------|
| Danneels, E. (2008). Organizational antecedents of second-order competences. <i>Strategic Management Journal</i> , 29(5), 519-543. | Auxiliary Diagnostics | No | No | No | |
| Semadeni, M., Cannella Jr, A. A., Fraser, D. R., & Lee, D. S. (2008). Fight or flight: Managing stigma in executive careers. <i>Strategic Management Journal</i> , 29(5), 557-567. | Heckman 2-stage | Yes | Yes | Not Reported | Evident but not discussed |
| Boyd, J. L., & Bresser, R. K. (2008). Performance implications of delayed competitive responses: Evidence from the US retail industry. <i>Strategic Management Journal</i> , 29(10), 1077-1096. | Diagnostics comparing sample with population | N/A | N/A | N/A | |
| Mitsuhashi, H., Shane, S., & Sine, W. D. (2008). Organization governance form in franchising: Efficient contracting or organizational momentum? <i>Strategic Management Journal</i> , 29(10), 1127-1136. | Lee's generalization of Heckman | No | Yes | Yes | |
| Mackey, A. (2008). The effect of CEOs on firm performance. <i>Strategic Management Journal</i> , 29(12), 1357-1367. | Auxiliary Heckman 2-stage | No | Not clear | Not Clear | |
| Meyer, K. E., Estrin, S., Bhaumik, S. K., & Peng, M. W. (2009). Institutions, resources, and entry strategies in emerging economies. <i>Strategic Management Journal</i> , 30(1), 61-80. | Auxiliary Heckman 2-stage | No | Not clear | No | |
| Yang, H., Lin, Z. J., & Lin, Y. L. (2010). A multilevel framework of firm boundaries: Firm characteristics, dyadic differences, and network attributes. <i>Strategic Management Journal</i> , 31(3), 237-261. | Heckman 2-stage | Yes | Yes | Yes | |
| Surroca, J., Tribó, J. A., & Waddock, S. (2010). Corporate responsibility and financial performance: The role of intangible resources. <i>Strategic Management Journal</i> , 31(5), 463-490. | Robust over Sub-samples | N/A | N/A | N/A | |
| Nadkarni, S., Herrmann, P., & Perez, P. D. (2011). Domestic mindsets and early international performance: The moderating effect of global industry conditions. <i>Strategic Management Journal</i> , 32(5), 510-531. | Auxiliary Heckman 2-stage | No | Not Clear | No | |
| Kotha, R., Zheng, Y., & George, G. (2011). Entry into new niches: The effects of firm age and the expansion of technological capabilities on innovative output and impact. <i>Strategic Management Journal</i> , 32(9), 1011-1024. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear as not reported | |
| Ethiraj, S. K., Ramasubbu, N., & Krishnan, M. S. (2012). Does complexity deter customer-focus? <i>Strategic Management Journal</i> , 33(2), 137-161. | Heckman 2-stage | No | Yes | Yes | |

| | | | | | |
|---|--|-----|-----------|---------------------------|--------------------------------|
| Diestre, L., & Rajagopalan, N. (2012). Are all 'sharks' dangerous? New biotechnology ventures and partner selection in R&D alliances. <i>Strategic Management Journal</i> , 33(10), 1115-1134. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear as not reported | |
| Lavie, D., Haunschild, P. R., & Khanna, P. (2012). Organizational differences, relational mechanisms, and alliance performance. <i>Strategic Management Journal</i> , 33(13), 1453-1479. | Inclusion as controls for variables which predict sample selection | No | No | N/A | Also evident but not discussed |
| Kang, J. (2013). The relationship between corporate diversification and corporate social performance. <i>Strategic Management Journal</i> , 34(1), 94-109. | Heckman 2-stage | Yes | Yes | No | |
| Wiersema, M. F., & Zhang, Y. A. (2013). Executive turnover in the stock option backdating wave: The impact of social context. <i>Strategic Management Journal</i> , 34(5), 590-609. | Heckman 2-stage | No | Yes | No | |
| Zhu, D. H. (2013). Group polarization on corporate boards: Theory and evidence on board decisions about acquisition premiums. <i>Strategic Management Journal</i> , 34(7), 800-822. | Heckman 2-stage | No | Yes | Yes | |
| Ndofor, H. A., Vanevenhoven, J., & Barker, V. L. (2013). Software firm turnarounds in the 1990s: An analysis of reversing decline in a growing, dynamic industry. <i>Strategic Management Journal</i> , 34(9), 1123-1133. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear | Also evident but not discussed |
| Mudambi, R., & Swift, T. (2014). Knowing when to leap: Transitioning between exploitative and explorative R&D. <i>Strategic Management Journal</i> , 35(1), 126-145. | Auxiliary Heckman 2-stage | No | Yes | Not Clear | |
| Eggers, J. P. (2014). Competing technologies and industry evolution: The benefits of making mistakes in the flat panel display industry. <i>Strategic Management Journal</i> , 35(2), 159-178. | Heckman 2-stage | Yes | Yes | No | |
| Xie, X., & O'Neill, H. M. (2014). Learning and product entry: How diversification patterns differ over firm age and knowledge domains in US generic drug industry. <i>Strategic Management Journal</i> , 35(3), 440-449. | Heckman 2-stage | No | Not Clear | Yes | |
| Stern, I., Dukerich, J. M., & Zajac, E. (2014). Unmixed signals: How reputation and status affect alliance formation. <i>Strategic Management</i> | Heckman 2-stage | Yes | No | No | |

| | | | | | |
|---|---------------------------|-----|-----|---------------------------|--|
| <i>Journal</i> , 35(4), 512-531. | | | | | |
| Krause, R., & Semadeni, M. (2014). Last dance or second chance? Firm performance, CEO career horizon, and the separation of board leadership roles. <i>Strategic Management Journal</i> , 35(6), 808-825. | Heckman 2-stage | No | Yes | No | |
| Berry, H. (2014). Global integration and innovation: Multicountry knowledge generation within MNCs. <i>Strategic Management Journal</i> , 35(6), 869-890. | 2-Stage Bivariate Probit | No | Yes | Yes | |
| Marcel, J. J., & Cowen, A. P. (2014). Cleaning house or jumping ship? Understanding board upheaval following financial fraud. <i>Strategic Management Journal</i> , 35(6), 926-937. | Auxiliary Heckman 2-stage | No | Yes | Not Clear as not reported | |
| Feldman, E. R., Gilson, S. C., & Villalonga, B. (2014). Do analysts add value when they most can? Evidence from corporate spin-offs. <i>Strategic Management Journal</i> , 35(10), 1446-1463. | Heckman 2-stage | Yes | Yes | Yes | |
| Patel, P. C., & Cooper, D. (2014). The harder they fall, the faster they rise: Approach and avoidance focus in narcissistic CEOs. <i>Strategic Management Journal</i> , 35(10), 1528-1540. | Heckman 2-stage | No | Yes | Yes | |
| Stettner, U., & Lavie, D. (2014). Ambidexterity under scrutiny: Exploration and exploitation via internal organization, alliances, and acquisitions. <i>Strategic Management Journal</i> , 35(13), 1903-1929. | Heckman 2-stage | No | Yes | Yes | |

Appendix B. Self-Selection Studies in SMJ (2002-2014).

| Article | Correction Method | Reported Selection Equation | Identification Variable(s) | Significant Selection | Sample Selection Present |
|---|---|-----------------------------|----------------------------|-----------------------|--------------------------|
| Nickerson, J. A., Hamilton, B. H., & Wada, T. (2001). Market position, resource profile, and governance: Linking Porter and Williamson in the context of international courier and small package services in Japan. <i>Strategic Management Journal</i> , 22(3), 251-273. | Instruments in a Recursive System | N/A | Yes | N/A | |
| Kale, P., Dyer, J. H., & Singh, H. (2002). Alliance capability, stock market response, and long-term alliance success: The role of the alliance function. <i>Strategic Management Journal</i> , 23(8), 747-767. | Heckman 2-stage | No | Not clear as not reported | Yes | |
| Leiblein, M. J., Reuer, J. J., & Dalsace, F. (2002). Do make or buy decisions matter? The influence of organizational governance on technological performance. <i>Strategic Management Journal</i> , 23(9), 817-833. | Heckman 2-stage & Lee procedures | Yes | Yes | Yes | |
| Leiblein, M. J., & Miller, D. J. (2003). An empirical examination of transaction-and firm-level influences on the vertical boundaries of the firm. <i>Strategic Management Journal</i> , 24(9), 839-859. | Tests for selection via Heckman procedure | No | Not Clear | No | |
| Brouthers, K. D., Brouthers, L. E., & Werner, S. (2003). Transaction cost-enhanced entry mode choices and firm performance. <i>Strategic Management Journal</i> , 24(12), 1239-1248. | Heckman 2-stage | Yes | Yes | Yes | |
| Krishnan, R. A., Joshi, S., & Krishnan, H. (2004). The influence of mergers on firms' product-mix strategies. <i>Strategic Management Journal</i> , 25(6), 587-611. | Heckman 2-stage | No | No | No | |
| Singh, K., & Mitchell, W. (2005). Growth dynamics: The bidirectional relationship between interfirm collaboration and business sales in entrant and incumbent alliances. <i>Strategic Management Journal</i> , 26(6), 497-521. | Heckman 2-stage | Yes | Yes | Yes | |
| Cantwell, J., & Mudambi, R. (2005). MNE competence-creating subsidiary mandates. <i>Strategic Management Journal</i> , 26(12), 1109-1128. | Heckman 2-stage & Lee procedures | Yes | No | Yes | |
| Miller, D. J. (2006). Technological diversity, related diversification, and firm performance. <i>Strategic Management Journal</i> , 27(7), 601-619. | Heckman's 2-stage | Yes | Yes | Yes | |

| | | | | | |
|---|---|-----|---------------------------|--------------|----------------------------|
| Arend, R. J. (2006). SME–supplier alliance activity in manufacturing: Contingent benefits and perceptions. <i>Strategic Management Journal</i> , 27(8), 741-763. | Heckman 2-stage | Yes | Yes | Yes | |
| Morrow, J. L., Sirmon, D. G., Hitt, M. A., & Holcomb, T. R. (2007). Creating value in the face of declining performance: Firm strategies and organizational recovery. <i>Strategic Management Journal</i> , 28(3), 271-283. | Heckman 2-stage | Yes | Yes | Yes | |
| Lazzarini, S. G. (2007). The impact of membership in competing alliance constellations: Evidence on the operational performance of global airlines. <i>Strategic Management Journal</i> , 28(4), 345-367. | Auxiliary Heckman 2-stage | No | Yes | No | |
| Ethiraj, S. K. (2007). Allocation of inventive effort in complex product systems. <i>Strategic Management Journal</i> , 28(6), 563-584. | Heckman 2-stage | Yes | Yes | Yes | |
| Williams, C. (2007). Transfer in context: Replication and adaptation in knowledge transfer relationships. <i>Strategic Management Journal</i> , 28(9), 867-889. | Structural Equation Modelling | N/A | N/A | N/A | |
| Capron, L., & Shen, J. C. (2007). Acquisitions of private vs. public firms: Private information, target selection, and acquirer returns. <i>Strategic Management Journal</i> , 28(9), 891-911. | Heckman 2-stage | Yes | No | No | |
| Lavie, D. (2007). Alliance portfolios and firm performance: A study of value creation and appropriation in the US software industry. <i>Strategic Management Journal</i> , 28(12), 1187-1212. | Auxiliary Heckman 2-stage | No | Not clear as not reported | Not Reported | |
| Sinha, R. K., & Noble, C. H. (2008). The adoption of radical manufacturing technologies and firm survival. <i>Strategic Management Journal</i> , 29(9), 943-962. | Heckman 2-stage | No | Not Clear | No | Also evident and discussed |
| Marcel, J. J. (2009). Why top management team characteristics matter when employing a chief operating officer: A strategic contingency perspective. <i>Strategic Management Journal</i> , 30(6), 647-658. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear | |
| Iyengar, R. J., & Zampelli, E. M. (2009). Self-selection, endogeneity, and the relationship between CEO duality and firm performance. <i>Strategic Management Journal</i> , 30(10), 1092-1112. | Switching Regressions Model | Yes | Yes | Yes | |
| Shamsie, J., Martin, X., & Miller, D. (2009). In with the old, in with the new: Capabilities, strategies, and performance among the Hollywood studios. <i>Strategic Management Journal</i> , 30(13), 1440-1452. | Auxiliary Heckman 2-stage and IV approaches | No | Not Clear | Not Reported | |

| | | | | | |
|---|--|-----|----------------------------|--------------|-------------------------|
| Zhang, Y., & Li, H. (2010). Innovation search of new ventures in a technology cluster: The role of ties with service intermediaries. <i>Strategic Management Journal</i> , 31(1), 88-109. | Heckman 2-stage | No | 3 Industry Dummy Variables | No | |
| Corredoira, R. A., & Rosenkopf, L. (2010). Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. <i>Strategic Management Journal</i> , 31(2), 159-181. | Auxiliary IV approach | No | Not Clear | N/A | |
| Chen, S. F. S. (2010). Transaction cost implication of private branding and empirical evidence. <i>Strategic Management Journal</i> , 31(4), 371-389. | Heckman 2-stage & Lee procedures | Yes | No | Yes | |
| Hoang, H. A., & Rothaermel, F. T. (2010). Leveraging internal and external experience: Exploration, exploitation, and R&D project performance. <i>Strategic Management Journal</i> , 31(7), 734-758. | Heckman's 2-stage with multinomial logit first stage | No | Yes | No | Also evident |
| Arikan, A. M., & Capron, L. (2010). Do newly public acquirers benefit or suffer from their pre-IPO affiliations with underwriters and VCs? <i>Strategic Management Journal</i> , 31(12), 1257-1289. | Heckman 2-stage | Yes | Yes | No | Also evident and tested |
| Parmigiani, A., & Holloway, S. S. (2011). Actions speak louder than modes: Antecedents and implications of parent implementation capabilities on business unit performance. <i>Strategic Management Journal</i> , 32(5), 457-485. | Heckman 2-stage | Yes | Yes | Yes | |
| Ho, J. L., Wu, A., & Xu, S. X. (2011). Corporate governance and returns on information technology investment: Evidence from an emerging market. <i>Strategic Management Journal</i> , 32(6), 595-623. | Auxiliary Heckman 2-stage & IV approaches | No | No | No | |
| Oh, C. H., & Oetzel, J. (2011). Multinationals' response to major disasters: How does subsidiary investment vary in response to the type of disaster and the quality of country governance? <i>Strategic Management Journal</i> , 32(6), 658-681. | Auxiliary Heckman 2-stage | No | Not Clear | Not Reported | |
| Gore, A. K., Matsunaga, S., & Eric Yeung, P. (2011). The role of technical expertise in firm governance structure: Evidence from chief financial officer contractual incentives. <i>Strategic Management Journal</i> , 32(7), 771-786. | Heckman 2-stage | Yes | Yes | Yes | |
| Hasan, I., Kobeissi, N., & Wang, H. (2011). Global equity offerings, corporate valuation, and subsequent international diversification. <i>Strategic Management Journal</i> , 32(7), 787-796. | Lee's generalization of Heckman | No | Yes | Yes | |

| | | | | | |
|--|---|-----|-----------|---------------------------|---|
| Jensen, P. H., Thomson, R., & Yong, J. (2011). Estimating the patent premium: Evidence from the Australian Inventor Survey. <i>Strategic Management Journal</i> , 32(10), 1128-1138. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear as not reported | |
| Park, H. D., & Steensma, H. K. (2012). When does corporate venture capital add value for new ventures? <i>Strategic Management Journal</i> , 33(1), 1-22. | IV approach via Bivariate probit model | Yes | Yes | N/A | |
| Quigley, T. J., & Hambrick, D. C. (2012). When the former CEO stays on as board chair: Effects on successor discretion, strategic change, and performance. <i>Strategic Management Journal</i> , 33(7), 834-859. | Heckman 2-stage | No | Yes | No | |
| Macher, J. T., & Boerner, C. (2012). Technological development at the boundaries of the firm: A knowledge-based examination in drug development. <i>Strategic Management Journal</i> , 33(9), 1016-1036. | Heckman 2-stage | Yes | Yes | Yes | |
| Jansen, J. J., Simsek, Z., & Cao, Q. (2012). Ambidexterity and performance in multiunit contexts: Cross-level moderating effects of structural and resource attributes. <i>Strategic Management Journal</i> , 33(11), 1286-1303. | Auxiliary Regressions fail to explain selection, and results robust over sub-samples. | No | Not Clear | N/A | |
| Weigelt, C. (2013). Leveraging supplier capabilities: The role of locus of capability deployment. <i>Strategic Management Journal</i> , 34(1), 1-21. | Lee's generalization of Heckman | Yes | Yes | Yes | Also evident and auxiliary Heckman employed to rule out |
| Lim, D. S., Celly, N., Morse, E. A., & Rowe, W. G. (2013). Rethinking the effectiveness of asset and cost retrenchment: The contingency effects of a firm's rent creation mechanism. <i>Strategic Management Journal</i> , 34(1), 42-61. | Heckman 2-stage | Yes | Yes | No | |
| Chung, C. N., & Luo, X. R. (2013). Leadership succession and firm performance in an emerging economy: Successor origin, relational embeddedness, and legitimacy. <i>Strategic Management Journal</i> , 34(3), 338-357. | Heckman 2-stage | No | Yes | Yes | |
| Mulotte, L., Dussauge, P., & Mitchell, W. (2013). Does pre-entry licensing undermine the performance of subsequent independent | Heckman 2-stage with ordered probit | Yes | Yes | Yes | |

| | | | | | |
|--|--|-----|-----------|-----|--|
| activities? Evidence from the global aerospace industry, 1944-2000. <i>Strategic Management Journal</i> , 34(3), 358-372. | | | | | |
| Miller, D., Minichilli, A., & Corbetta, G. (2013). Is family leadership always beneficial? <i>Strategic Management Journal</i> , 34(5), 553-571. | Heckman 2-stage | No | Yes | Yes | |
| Zahavi, T., & Lavie, D. (2013). Intra-industry diversification and firm performance. <i>Strategic Management Journal</i> , 34(8), 978-998. | Heckman 2-stage | Yes | No | No | |
| Lahiri, N., & Narayanan, S. (2013). Vertical integration, innovation, and alliance portfolio size: Implications for firm performance. <i>Strategic Management Journal</i> , 34(9), 1042-1064. | Heckman 2-stage | Yes | Yes | Yes | |
| Carson, S. J., & John, G. (2013). A theoretical and empirical investigation of property rights sharing in outsourced research, development, and engineering relationships. <i>Strategic Management Journal</i> , 34(9), 1065-1085. | Procedure where residuals from selection equation are entered as regressors in the substantive equation. | Yes | Yes | Yes | |
| Tong, T. W., & Li, S. (2013). The assignment of call option rights between partners in international joint ventures. <i>Strategic Management Journal</i> , 34(10), 1232-1243. | Heckman 2-stage | No | Not Clear | No | |
| Wu, B. (2013). Opportunity costs, industry dynamics, and corporate diversification: Evidence from the cardiovascular medical device industry, 1976–2004. <i>Strategic Management Journal</i> , 34(11), 1265-1287. | 2-step procedure | Yes | Yes | Yes | |
| Yang, H., Zheng, Y., & Zhao, X. (2014). Exploration or exploitation? Small firms' alliance strategies with large firms. <i>Strategic Management Journal</i> , 35(1), 146-157. | Heckman 2-stage | No | No | No | |
| Jia, N. (2014). Are collective political actions and private political actions substitutes or complements? Empirical evidence from China's private sector. <i>Strategic Management Journal</i> , 35(2), 292-315. | Auxiliary 2-stage probit model | Yes | Yes | N/A | |
| Patel, P. C., Fernhaber, S. A., McDougall-Covin, P. P., & van der Have, R. P. (2014). Beating competitors to international markets: The value of geographically balanced networks for innovation. <i>Strategic</i> | Heckman 2-stage | No | Yes | Yes | |

| | | | | | |
|---|--|-----|-----------|---------------------------|--------------------------------|
| <i>Management Journal</i> , 35(5), 691-711. | | | | | |
| Laamanen, T., Brauer, M., & Junna, O. (2014). Performance of acquirers of divested assets: Evidence from the US software industry. <i>Strategic Management Journal</i> , 35(6), 914-925. | Heckman 2-stage with multivariate choices | No | Not Clear | No | |
| Pathak, S., Hoskisson, R. E., & Johnson, R. A. (2014). Settling up in CEO compensation: The impact of divestiture intensity and contextual factors in refocusing firms. <i>Strategic Management Journal</i> , 35(8), 1124-1143. | Auxiliary Heckman 2-stage | No | Yes | No | |
| Heide, J. B., Kumar, A., & Wathne, K. H. (2014). Concurrent sourcing, governance mechanisms, and performance outcomes in industrial value chains. <i>Strategic Management Journal</i> , 35(8), 1164-1185. | Auxiliary Heckman 2-stage | No | Yes | No | Also evident but not discussed |
| Herrmann, P., & Nadkarni, S. (2014). Managing strategic change: The duality of CEO personality. <i>Strategic Management Journal</i> , 35(9), 1318-1342. | Auxiliary Heckman 2-stage | Yes | Yes | No | |
| Jain, A., & Thiétart, R. A. (2014). Capabilities as shift parameters for the outsourcing decision. <i>Strategic Management Journal</i> , 35(12), 1881-1890. | Auxiliary Heckman 2-stage and Switching Regressions approaches | No | No | No | |
| Shen, R., Tang, Y., & Chen, G. (2014). When the role fits: How firm status differentials affect corporate takeovers. <i>Strategic Management Journal</i> , 35(13), 2012-2030. | Auxiliary Heckman 2-stage | No | Not Clear | Not Clear as not reported | |
| Rubera, G., & Tellis, G. J. (2014). Spinoffs versus buyouts: Profitability of alternate routes for commercializing innovations. <i>Strategic Management Journal</i> , 35(13), 2043-2052. | Auxiliary Heckman 2-stage | No | Yes | No | |