RESEARCH ARTICLE

GLOBAL STRATEGY JOURNAL   WILEY

Tobit models in strategy research: Critical issues and applications

Mario Daniele Amore[1] | Samuele Murtinu[2]

[1]Bocconi University, Milan, Italy

[2]University of Groningen, Groningen, the Netherlands

Correspondence
Samuele Murtinu, University of Groningen, Nettelbosje 2 - 9747 AE, Groningen, the Netherlands.
Email: s.murtinu@rug.nl

Abstract

Research Summary: Tobit models have been used to address several questions in management research. Reviewing existing practices and applications, we discuss three challenges: (a) assumptions about the nature of data, (b) apparent interchangeability between censoring and selection bias, and (c) potential violations of key assumptions in the distribution of residuals. Empirically analyzing the relationship between import competition and industry diversification, we contrast Tobit models with results from other estimators and show the conditions that make Tobit a suitable empirical approach. Finally, we offer suggestions and guidelines on how to use Tobit models to deal with censored data in strategy research.

Managerial Summary: Data on strategic decisions often exhibit certain features, such as excess zeros and values bounded within a given range, which complicate the use of linear econometric techniques. Deriving statistical evidence in such instances may suffer from biases that undermine managerial applications. Our study presents an extensive comparison of different econometric models to deal with censored data in strategic management showing the strengths and weaknesses of each model. We also conduct an application to the context of import penetration and industry diversification to highlight how the relationship between these two variables changes depending on the econometric model used for the analysis. In conclusion, we provide a set of recommendations for scholars interested in censored data.

1 | INTRODUCTION

Many strategic decisions are composed of two decisions: a "yes/no" choice about doing or not a certain activity and, in the case of a "yes," a "how much" choice about the amount of resources to dedicate to such an activity. These settings include, for instance, corporate ownership choices (e.g., Delios & Beamish, 1999), entry modes (e.g., Luo & Bu, 2018), diversification (e.g., Bowen & Wiersema, 2005; Reuer & Leiblein, 2000; Tan & Chintakananda, 2016), acquisitions (e.g., Devers, McNamara, Haleblian, & Yoder, 2013; Ragozzino & Reuer, 2011), and R&D/innovation strategies (e.g., Cassiman & Veugelers, 2006; Laursen & Salter, 2006; Sorenson, McEvily, Ren, & Roy, 2006). In all these settings, researchers often observe data on the "how much" decision but may need to make assumptions on the "yes/no" decision. Empirically, data about the "how much" decision may display a density mass at a certain value of its distribution—usually at zero. In these situations, researchers may confound the *observed* and the *latent* dependent variables in their regression analysis.

For example, consider a firm's strategic decision about engaging in foreign acquisitions (i.e., the dollar amount allocated to foreign acquisitions): a fully observed dependent variable means that the zeros are *true zeros* representing the actual choice of the firm under investigation (i.e., the firm could potentially engage in foreign acquisitions but chooses *not to*). In this case, the dependent variable is assumed to follow a mixed distribution where there is a probability mass at zero and a continuous distribution for values greater than zero. The zeros and the non-zero values thus come from the same data generating process (a data setting typically called corner solution). The nature of such dependent variable makes ordinary least squares (OLS, henceforth) potentially unsuitable, and Tobit models may constitute a valid estimation approach.[1] Further, as zeros are true zeros (and not, for instance, imputed values to missing data), there is no selection bias.

By contrast, when (some of) the zeros are not *true zeros*, the dependent variable is only partially observed. Here, we can have two different situations. The first is when the non-zero values are true observations of the dependent variable, but the zeros indicate that the value of the dependent variable is missing (i.e., latent). Here, there may be a selection bias because the zeros and the non-zeros may arise from two distinct stochastic processes; as we will show, Tobit and Heckman models produce different results. The second situation relates to data censoring. For example, values about the dollar amount of foreign acquisitions below a certain threshold could have been coded as zero. In this case, it is impossible to distinguish between the *true* zeros (firms that do not engage in foreign acquisitions at all) and the *false* zeros (firms that invest little in foreign acquisitions). Notice that while in the case of corner solutions we have a kink in the distribution of the variable of interest, data censoring may arise from a problem of data observability (not associated with selection bias). In these situations, we will show that the use of Tobit is appropriate.

As we argue in this study, even in the cases where Tobit models are suitable, their use and interpretation are often complicated by a number of critical issues. Conducting a comprehensive review of studies published in five top management journals from 1980 to 2015, we found that around 47% of the articles suffer from potentially misleading assumptions about the nature of zeros. Moreover,

we identified two other critical issues: the idea that Tobit models could address problems of selection bias, and potential violations of Normality and homoscedasticity in the distribution of residuals.

As regards the former, Tobit models assume that the variables explaining whether or not the *observed* dependent variable is censored must also explain the *level* of the variable when it takes positive values. Given that this assumption may not hold in samples affected by selection bias, or when the "yes/no" choice and the "how much" choice are explained by different mechanisms, the use of Tobit models may lead to unreliable estimates. In our review of the literature, this issue appeared in 7% of the studies if we only include studies where the authors explicitly state that Tobit models are used to address selection bias. If we also consider studies where the authors implicitly argue that Tobit models are the most suitable choice, the issue is far more common. Finally, neglecting potential violations of Normality and homoscedasticity in the distribution of residuals can be problematic since Tobit models crucially hinge on these assumptions. Especially in small samples, violations of these assumptions lead to unreliable inference. We find that almost 53% of the reviewed studies do not explicitly account for these issues.

We test the importance of the three above issues by comparing Tobit with OLS, Heckman, and two-part models (namely, the Truncated Normal Hurdle model developed by Cragg, 1971). To this end, we use firm-level data to revisit extant evidence on the relationship between import competition and industry diversification among US companies. First, we document that when zeros are true zeros in a corner solution setting, Tobit models are better (less) suited than OLS (two-part models).[2] Second, we show that Tobit models are not interchangeable with Heckman models in addressing selection bias. Finally, we show that if residuals are wrongly assumed as homoscedastic, there will be an over-rejection of the null hypothesis. We provide further evidence on these issues as well as on the role played by sample size through Monte Carlo simulated data (available in the Appendix). In so doing, we significantly expand existing efforts to understand the appropriateness of Tobit models (Bowen & Wiersema, 2004) as well as, more generally, the ongoing methodological debates in strategy research (e.g., Certo, Busenbark, Woo, & Semadeni, 2016; Semadeni, Withers, & Certo, 2014).

After a discussion of the main methodological issues behind the use of Tobit models, we provide a comprehensive set of guidelines that strategy scholars may follow when dealing with censored data and Tobit models in their empirical studies. We wish that this work will improve the familiarity of researchers with the use and interpretation of censored data in strategy research.

2 | OVERVIEW OF TOBIT MODELS

2.1 | Basic framework

Tobit models (Tobin, 1958) belong to a class of econometric techniques traditionally regarded as censored regression models (Wooldridge, 2002). To start, it is worth clarifying the difference between censoring, truncation, and corner solutions. Censoring is sometimes present in datasets containing information on R&D investment where data providers may recode all values of R&D intensity (e.g., R&D scaled by revenues) above a given threshold with such threshold value to avoid the identification of single firms. For instance, Becker and Dietz (2004) have an upper (or right) censoring in which values of R&D intensity above 0.35 are set equal to 0.35.[3] As an example of truncation suppose, for instance, to assess the impact of certain individual characteristics (e.g., gender) on labor income; however, due to privacy concerns of the data provider, income is not observable for the whole population but only for those individuals who earn more than $20,000 per year. In this case, the dependent variable (i.e., income) is truncated: values below $20,000 are coded as missing. Corner solutions represent cases

where a given variable, say consumption, is observable and exhibits a continuous distribution over positive values but it is equal to zero for some individuals as a result of an optimization problem.

Having clarified these concepts, we can define a Tobit model as follows:

$$y^* = X'\beta + \varepsilon, \text{ with } \varepsilon \,|\, X \sim N(0;\sigma^2), \text{ with } y = y^* \text{if } y^* > 0, \text{ and } y = 0 \text{ otherwise.} \quad (1)$$

where y is the observed variable of interest, and y^* is the latent variable. Equation (1) states three things. First, the expected *effect* of X on y^* is monotonic. Second, the residuals follow a Normal distribution. Third, the dependent variable is *left-censored*.[4]

In Table 1, we show an application of Tobit models using a Monte Carlo method to build a sample of 100 observations where y^*, y and X follow Equation (1).[5] In Panel A, we focus on the linear part of the model, that is, we only use observations with y greater than zero (*uncensored* observations). Comparing OLS and Tobit estimates, we can see that the coefficients are the same. In Panel B, we estimate the full model, that is, including also observations with y equal to 0. As shown, the coefficient of X estimated with OLS is lower than the one estimated with Tobit; in this specific application, ignoring censoring in OLS translates into a lower slope of the regression line and an inflated intercept.
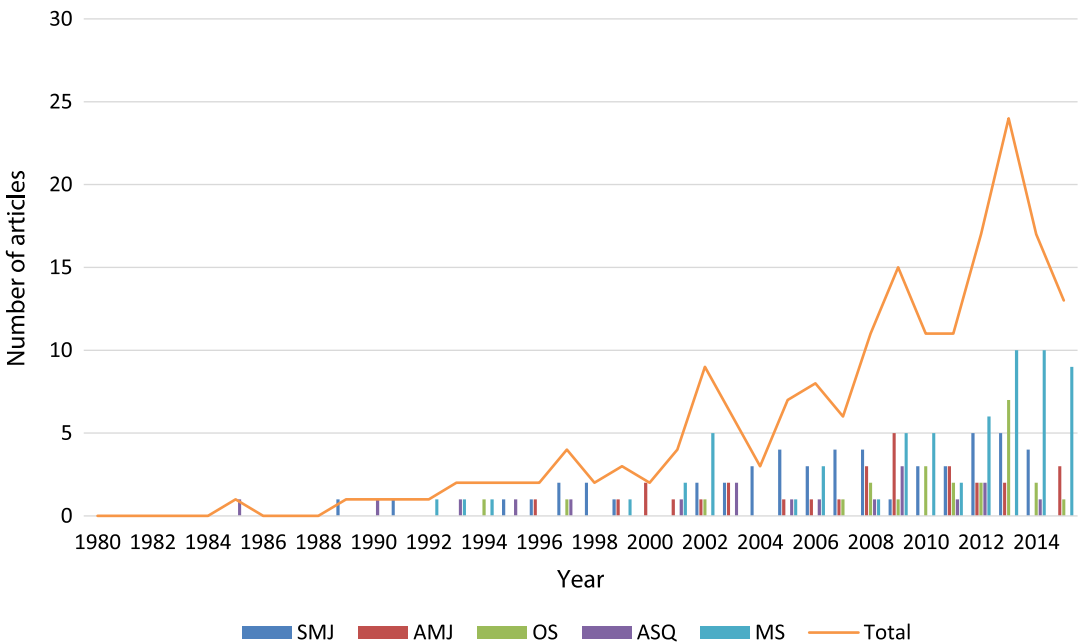
2.2 | Tobit models in management research

To identify empirical applications of Tobit models in management research, we conducted a full-text search for the keyword "Tobit" in its various forms (e.g., "Tobit model(s)," "Tobit regression(s)") in all articles published in the Academy of Management Journal (AMJ), Administrative Science Quarterly (ASQ), Management Science (MS), Organization Science (OS), and Strategic Management Journal (SMJ) from 1980 to 2015. Then, we augmented our search with other relevant keywords such as "censoring" and "truncation."

TABLE 1 Comparison between OLS and Tobit

	OLS	Tobit
Panel A		
Constant	1.9832	1.9832
	(.000)	(.000)
X	0.9073	0.9073
	(.000)	(.000)
Observations	85	85
Censored obs.	0	0
Panel B		
Constant	1.8600	1.6984
	(.000)	(.000)
X	0.9690	1.1598
	(.000)	(.000)
Observations	100	100
Censored obs.	15	15

Note: Two-sided *p*-values are reported in parentheses.

FIGURE 1 Tobit models in management research

After a manual screening of each article to avoid double counting and keep only articles that empirically estimate a Tobit model (and not just refer to applications of Tobit models used elsewhere), we found a total of 186 articles, of which 29 are in AMJ, 18 in ASQ, 63 in MS, 24 in OS, and 52 in SMJ. To avoid arbitrariness, each of the authors went through the articles independently. Figure 1 shows an upward trend in the use of Tobit models—especially driven by MS in the 2012–2015 period.[6]

Analyzing these articles, we focused on identifying three main issues that may complicate the use and interpretation of Tobit models, namely: (a) potentially wrong assumptions about the nature of the data; (b) apparent interchangeability between censoring and selection bias; and (c) disregard of potential violations of Normality and homoscedasticity in the distribution of residuals. These three insidious features cover the most fundamental aspects that can be commonly misinterpreted in regression methods: the nature of the data used to build the dependent variable, the specification of the regression model, and the structure of residuals. Complementing a wide econometric literature on these topics (e.g., Arabmazar & Schmidt, 1981, 1982; Bowen & Wiersema, 2004), in the next section we conduct an empirical study of a global strategy question to illustrate the main challenges of using Tobit models. We then provide a set of guidelines to scholars who want to use Tobit models.

3 | AN APPLICATION WITH FIRM-LEVEL DATA

A long-running literature in global strategy has sought to estimate the effect of foreign competition on corporate diversification strategies. An ideal setting to address this question is provided by the reduction in tariff barriers, which leads to an increase in competitive pressures due to stronger foreign competition (e.g., Bowen & Wiersema, 2005). Revisiting the existing evidence on this topic, we investigate the role played by the nature of the dependent variable, the way to tackle sample selection

issues, and the assumptions made on the distribution of residuals. In the Appendix, we provide further evidence on these issues using simulated data.

3.1 | Sample and variables

We use the sample of US listed firms covered in the Compustat dataset starting from 1976.[7] To measure a firm's level of industry diversification, we use the historical segment data (containing sales by geographic areas and industries) to compute the Herfindahl–Hirschman Index (HHI) of the concentration of a firm's revenues across 4-digit SIC industries. We take one minus HHI, such that greater values correspond to higher levels of firm diversification. The resulting variable (*Diversification*) is bounded within the [0, 1] range and will be used as *dependent variable* in our analysis. In the final sample, 43% of observations correspond to undiversified firms (i.e., firms for which the dependent variable is zero).

Our main *explanatory variable* (*Import penetration*) comes from Peter Schott's archive (see also Bernard, Jensen, & Schott, 2006) and is computed as the 1-year lagged imports divided by domestic absorption (i.e., the sum of gross investment, and household and government consumption) for each 4-digit SIC manufacturing industry (i.e., from 2,000 to 3,999) and each year until 1999.

We then build a set of *control variables* related to a firm's size and financial conditions, as well as to the industry's attractiveness, concentration, and innovativeness. In particular, we use Compustat to compute a firm's return on assets (ROA) as the ratio of earnings before interest, taxes, depreciation and amortization (EBITDA) to total assets, and the logarithm of a firm's sales value. Moving to the industry level, we use Compustat data to compute the following variables at the 4-digit SIC and year level: industry ROA; core business profitability, measured as the ratio of operating profits to revenues; industry concentration, computed as the HHI of revenues; and industry R&D intensity, computed as the ratio of R&D expenditures to sales.[8] From the NBER manufacturing dataset we obtain a measure of industry capital intensity, that is, the ratio of real capital stock to total employment. Finally, we augment our model with year dummies to control for time effects common to all firms, and 4-digit SIC industry dummies to control for time-invariant heterogeneity across industries.

After dropping observations with missing values, we obtain a final sample of 4,857 unique firms and 40,153 observations from 1976 to 1999.

3.2 | Empirical approach

Motivating their choice of a Tobit model, Bowen and Wiersema (2005) write that: "Almost 60 percent of the 8,961 observations in our dataset are single business firms whose level of diversification—our dependent variable—has a calculated value of zero. When a high proportion of the values taken by a dependent variable equals a single 'limit value' (here zero), an appropriate estimation technique is the nonlinear Tobit procedure" (p. 1161).

In Panel A of Table 2 we compare the estimates obtained from OLS, Tobit,[9] and two-part models. As an example of this latter class of models, we use the Truncated Normal Hurdle (TNH) model developed by Cragg (1971).[10] In the case of a corner solution (as the application here considered) we cannot directly compare the estimated coefficients. In fact, after Tobit regressions we can derive various marginal effects depending on whether we are interested in the effect on the expected value of the latent variable or on the unconditional expected value of the observed variable.[11] To get the marginal effect of a regressor (X) on the observed variable we must multiply the coefficient β by the probability that the observed variable is greater than zero, that is ($X'\beta/\sigma$): $\delta E(y)/\delta X = \beta * \Phi(X'\beta/\sigma)$. In

TABLE 2 An application with firm-level data

	Full sample			Only diversified firms		
	OLS	**Tobit**	**TNH**	**OLS**	**Tobit**	**TNH**
Import penetration	−0.0156	−0.0290	−0.0837	−0.0527	−0.0527	−0.0527
	(.292)	(.028)	(.000)	(.010)	(.010)	(.010)
Drukker's test		101.59 (.000)			419.73 (.000)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	40,153	40,153	40,153	23,026	23,026	23,026

Panel A. Nature of data censoring appears at the top of the table.

Panel B. Data censoring versus sample selection

				Heckman		Tobit
Import penetration				−0.0488		−0.0290
				(.018)		(.028)
Controls				Yes		Yes
Industry fixed effects				Yes		Yes
Year fixed effects				Yes		Yes
Observations				40,153		40,153

Panel C. SEs in Tobit models

	Unadjusted			Clustered by industry		
Import penetration	−0.0290			−0.0290		
	(.028)			(.244)		
Controls	Yes			Yes		
Industry fixed effects	Yes			Yes		
Year fixed effects	Yes			Yes		
Clusters	122			122		
Observations	40,153			40,153		

Note: Two-sided *p*-values are reported in parentheses. For the Heckman model in Panel B, the corresponding inverse Mill's ratio is −0.0479 (*p*-value = .009).

all Panels of Table 2, we report the marginal effects on the observed variable. Specifically, instead of calculating the marginal effect of *X* at the mean value of the regressor, we estimate average marginal effects (i.e., the average of partial changes over all observations).

We carry out the estimations separately for two samples. The first is the full sample, where, as discussed above, *Diversification* contains 43% of zeros and 57% of strictly positive values. The second is the subsample of diversified firms (in which zeros are excluded). Our results show two different patterns. When we focus on the subsample of diversified firms (i.e., where the dependent variable does not contain zeros), we find that the marginal effects of OLS, Tobit and TNH have the same sign, magnitude, and statistical significance. By contrast, when we employ the full sample, the marginal effect of *Import penetration* is negative and statistically significant in Tobit and TNH

estimates, whereas it is insignificant in OLS estimates.[12] Assuming that the zeros in the distribution of the dependent variable are true zeros, as wisely done by Bowen and Wiersema (2005), this discrepancy can be attributable to the fact that OLS does not appropriately account for the high proportion of zeros.[13]

The difference in magnitude between the marginal effects estimated by means of Tobit and TNH models (with the latter being more than three times larger than the former) depends on the mechanisms governing the decision to diversify (or not) and the decision about the degree of diversification, that is, "whether the zero and the positive observations are generated by the same mechanism" (Silva, Tenreyro, & Windmeijer, 2015, p. 29). Whether these two decisions are potentially explained by two different mechanisms (as allowed by the TNH model) or by a single mechanism (as assumed by Tobit) is often an empirical question. To figure this out, one can estimate a TNH model[14] and check whether there is any covariate whose coefficient has a different sign in the first step vis-à-vis the second step of the TNH estimation. For instance, in our application the (unreported) coefficient of industry capital intensity is negative (and not statistically significant) in the first step and positive (and statistically significant at the 1% level) in the second step. This evidence runs against the assumption of Tobit models that the determinants of the binary decision must also explain—with the same sign—the intensity decision.[15] In this setting, the TNH model represents the most suitable approach. Indeed, differently from Tobit, in two-part models the association between the covariates and the decision to diversify can be different from the association between the same covariates and the degree of diversification. To this latter extent, another advantage of the TNH model is that it allows to have different covariates in the two steps (Cameron & Trivedi, 2010).

In Panel B of Table 2, we explore the implications that arise from confounding corner solutions (in this case, there is a corner at zero, and a continuous distribution bounded at one) with selection bias. We contrast estimates from Tobit and Heckman models, with this latter being the conventional approach to address selection issues. The selection issue we seek to solve with an Heckman model is the one happening between firms that diversify and firms that do not. As long as the zeros are not imputed values to missing data (as in our sample, which excludes observations with missing data in the variable *Diversification*), and thus *Diversification* is always observed, selection bias problems concerning the zeros are unlikely to be present. Our results show that Tobit and Heckman models yield different results with the absolute value of Tobit marginal effect being almost 1.7 times smaller than the Heckman estimate. Even though our empirical analysis employs a relatively large sample, the difference between the two marginal effects is quite large. To understand which method is more suitable, we use the statistical test for corner solutions proposed by Silva et al. (2015). The goal of this approach is to compare the conditional expectations of the dependent variable obtained by different models, that is, whether an estimate under an alternative model improves the prediction of the dependent variable obtained by means of the baseline model. Choosing Tobit as the baseline model and Heckman as the alternative model, we run the above test and we do not reject the null hypothesis (p-value = .99), that is, Tobit is valid.

Methodologically, it is important to notice that in order to estimate the selection equation of the Heckman model we did not employ any exclusion restriction (i.e., an additional explanatory variable that predicts the binary choice to diversify while not affecting how much to diversify). As Dow and Norton (2003) argue, "exclusion assumptions are often unavailable or hard to defend" (p. 9). In the case of corporate diversification, the search for variables that could correlate with the binary decision to diversify or not without explaining the intensive margin of how much to diversify is still unsettled after several decades of research. The difficulty is due to the fact that the two decisions are essentially set jointly (i.e., they are an equilibrium point arising from managing unobservable tradeoffs within

the firm). Any imprecise exclusion restriction will raise empirical concerns which can aggravate our estimation (Bound, Jaeger, & Baker, 1995). At the same time, estimations without exclusion restrictions can be problematic due to potential insufficient identifying variation to estimate the coefficient of interest (that of *Import penetration* in our case) in the main equation (Wolfolds & Siegel, 2019). Following Madden (2008), we mitigate this concern by verifying that the variance inflation factor associated to the inverse Mills' ratio (IMR)—the selection-correction term added to the main equation—is not above 10, the common threshold used in the literature to detect collinearity concerns.[16]

In Panel C of Table 2 we deal with issues concerning the distribution of residuals. Recall that Tobit models crucially rely on Normal and homoscedastic residuals—assumptions which are often violated in panel data settings like ours. We show the importance of accounting for the specific structure of residuals by comparing unadjusted residuals (i.e., assuming homoscedasticity) with residuals adjusted by clustering at the industry level. The rationale behind this choice is that standard errors (*SE*s) are likely heteroscedastic and serially correlated due to group (within cluster) correlation arising from industry characteristics. For instance, firm diversification choices may be driven by industry-level dynamics over time, in the form of technological shocks, changes in export and/or import competition, and foreign direct investments. In the presence of nested two-way clustering (for instance, firm-level and industry-level clustering), some scholars suggest to cluster *SE*s at the highest level of aggregation (Cameron, Gelbach, & Miller, 2011; Pepper, 2002). Thus, in our setting it may be advisable to cluster residuals at the industry level (for more discussion about the proper dimension of clustering see Section 4.3; see also Bertrand, Duflo, & Mullainathan, 2004; Petersen, 2009 and, more recently, Abadie, Athey, Imbens, & Wooldridge, 2017). As shown, once we adopt the industry clustering procedure, *SE*s become almost twice as large as the unadjusted ones, making the marginal effect of *Import penetration* not statistically different from zero.

Finally, it is worth noting that the empirical evidence on the relationship between foreign competition and corporate diversification may suffer from endogeneity issues due to unobserved heterogeneity at the firm level. Even though a thorough investigation of endogeneity issues in our empirical application is beyond the scope of the work, addressing endogeneity in presence of panel censored data is a relevant issue. Due to incidental parameter problems,[17] Tobit models in panel settings cannot be estimated by means of fixed effects estimation. Two alternatives are the semiparametric trimmed least absolute deviation (LAD) estimator with fixed effects (Honoré, 1992), and the panel data regression model with two-sided censoring (Alan, Honoré, Hu, & Leth-Petersen, 2014).

4 | SUMMARY AND RECOMMENDATIONS

In order to provide relevant business and policy implications, strategy researchers face the challenge of identifying precisely certain empirical relationships. Toward this end, scholars are making significant efforts to improve our understanding of different estimation approaches in the context of strategy research (e.g., Blevins, Tsang, & Spain, 2015; Certo et al., 2016; Clougherty, Duso, & Muck, 2016; Hamilton & Nickerson, 2003; Hoetker, 2007; Ketchen & Shook, 1996; Molina-Azorin, 2012; Peterson, Arregle, & Martin, 2012; Shook, Ketchen, Cycyota, & Crockett, 2003; Shook, Ketchen, Hult, & Kacmar, 2004; Wiersema & Bowen, 2009).

Tobit models are widely used to deal with censored dependent variables. Our review of scholarly work in leading management journals from 1980 to 2015 has detected a growing number of applications of Tobit models in several areas from strategy to organization and innovation management. Despite many advantages, Tobit models may lead to imprecise estimates when scholars are

misguided in discerning the nature of the dependent variable, the difference between selection concerns and censored data, and the distribution of the residuals. How could scholars avoid these problems? Existing methodological inquiries have assessed the use of limited dependent variable models in strategy research (Wiersema & Bowen, 2009); however, such inquiries have mostly focused on Logit and Probit models (Hoetker, 2007; Wiersema & Bowen, 2009) or on the strengths and weaknesses of Tobit by comparing it with OLS (Bowen & Wiersema, 2004; Mudambi & Helper, 1998). Our work provides an ideal complement to these existing efforts by guiding strategy scholars in the practical implementation of models featuring censoring, corner solutions, truncation and/or selection bias. Adding to the work by Bowen and Wiersema (2004) and Mudambi and Helper (1998), our analysis of censoring and selection bias compares Tobit with a broader set of alternative estimators (OLS, Heckman and TNH models), and empirically analyzes issues regarding the distribution of residuals in Tobit models. Our enquiry also provides easy-to-implement stepwise procedures to properly estimate data featuring censoring and selection bias. Collectively, our discussions provide guidance on what estimation scholars should do when dealing with censored or bounded dependent variables.

4.1 | Understanding the nature of the dependent variable

The first common pitfall in the use of Tobit models comes from potentially misleading interpretations of the dependent variable, which may not necessarily be censored even when it takes values within certain ranges or has a density mass at given points of its distribution. To figure the precise nature of their dependent variable, strategy scholars should address the following questions. Is the dependent variable censored, truncated or does it display a corner solution? If so, why does it display these features? To answer these questions, scholars need to think about the theoretical or empirical processes that create the censoring or corner solution, and/or the coding procedures put in place by data providers. Occasionally, coding procedures lead to truncation, that is, the dependent variable is extracted from a subset of the whole population. In these cases, Tobit models are not the most suitable choice, and scholars should opt for truncated regression models.

Once scholars have clearly understood the nature of data censoring, it is important to address the following question: What are the specific thresholds of censoring (which may be inferred from the data collection process or existing research)? If the dependent variable is an uncensored proportion (e.g., theoretically bounded between 0 and 100% without any censoring) scholars should consider the benefits of specific models such as the fractional Logit (see Papke & Wooldridge, 1996, 2008, Wulff & Villadsen, 2019 and Baum, 2008 for an application using the Stata package). Instead, if the dependent variable shows many zeros, and the researcher assumes that these zeros are *true zeros* representing the actual choice of the economic agents under investigation (e.g., firms could potentially engage in a diversification strategy but choose not to), Tobit models may represent a valid choice when the zeros and the positive observations are driven by the same mechanism.

4.2 | Accounting for selection versus censoring issues

The second pitfall arises from an apparent interchangeability between sample selection and data censoring/corner solutions. Examples are found in studies dealing with R&D expenses, corporate diversification or geographic distance in investment decisions (typically displaying several zeros). When used to address sample selection, Tobit and Heckman models produce different estimates.

Conceptually, scholars need to ask whether they are correctly distinguishing between sample selection from corner solutions or censoring. As regards corner solutions (assumed at zero), there is a density mass at zero; however, as long as the zeros are not imputed values to missing data, there is no selection problem to address concerning the zeros, and Tobit models may be an appropriate choice. By contrast, if zeros correspond to observations for which the dependent variable is missing, the researcher needs to test whether zeros and non-zeros systematically differ according to some characteristics (which are observable for both zeros and non-zeros). For instance, the researcher can conduct a series of *t* tests for the equality of means for all of the covariates through which zero and non-zero observations are compared—or, alternatively, a LR (likelihood ratio) test on the joint insignificance of mean covariate differences between zero and non-zero observations. Rejecting the null hypothesis in these tests may point to the presence of selection bias. More formally, this means that the assumption that the "yes/no" decision *dominates* the "how much" decision (i.e., the zeros in the selection equation come from a separate discrete decision rather than a corner solution) is likely to hold in the data, and thus Tobit models are not suitable (see Madden, 2008).

As Angrist (2001) argues, the choice between Heckman-type and two-part models[18]—Heckman models where the correlation between the selection equation and the main equation is assumed to be zero (and so there is no need to include the IMR term in the main equation) and where the residuals in the main equation are not necessarily Normally distributed—also depends on the interest of the researcher over the observed variable vis-à-vis the latent variable. In the former case, two-part models may be preferred because of less structural assumptions (practically, scholars can estimate two separate regressions: a Probit for the binary decision, and an OLS for the intensity decision on the sub-sample of non-zero observations). In the latter case, Heckman-type models are the most suitable choice.[19] Whichever the dependent variable of interest, if the concern is that of selection bias, then scholars should opt for models specifically designed to address selection issues.

A useful three-step approach to choose a suitable model to address selection bias is the following. First, estimate an Heckman model with a reliable exclusion restriction (i.e., an additional explanatory variable which predicts the binary selection variable while not affecting the dependent variable in the main equation). Second, run a LR test (which is often automatically implemented in statistical software packages) on the independence of the selection and the main equation. If the two equations are independent ($\rho = 0$)—the binary decision (e.g., to diversify or not) is not influenced by the intensity decision (e.g., how much to diversify)—the Heckman model should be abandoned. Third, one should ask whether the *binary* decision and the *intensity* decision are sequential or simultaneous (Humphreys, 2013; Jones, 2000). If the two decisions are sequential, then scholars should opt for two-part models (Aitchison, 1955; Cragg, 1971; Duan, Manning, Morris, & Newhouse, 1983; Farewell, Long, Tom, Yiu, & Su, 2017; Humphreys, 2013; Jones, 2000).

It is worth stressing that, even if the above LR test calls for the independence of the selection and the main equation, scholars need to reason on unobservable factors (not included in the model specification) that potentially affect both the selection and the main equation. Indeed, an assumption behind two-part models is that unobservable factors that influence the selection equation are uncorrelated with unobservable factors affecting the main equation. However, in the context of our research question it is not difficult to think about unobserved variables (i.e., excluded from our model specification), such as managerial foreign experience, which can affect both the decision to diversify and the one about how much to diversify.

Despite their many advantages (such as the ease of estimation, and minimal computational problems and distributional assumptions), two-part models have been designed to identify only the *observed* variable of interest (rather than the *latent* variable). Indeed, being the coefficients of

regressors in the linear part of the model estimated only for the non-zero values, it is challenging to calculate marginal effects on the whole distribution of values of the dependent variable (that can be instead easily computed after Tobit estimations).[20] Also for these reasons, the merits of two-part models are debated in the literature (see, for instance, Leung & Yu, 1996).

4.3 | Dealing correctly with the distribution of Tobit residuals

The third pitfall concerns the distribution of residuals. Residuals of Tobit estimations are often non-Normally distributed and/or heteroscedastic (and serially correlated in panel applications), and neglecting these features causes misleading *SE*s. Unfortunately, scholars cannot use standard Lagrange Multiplier (LM) tests for Normality and homoscedasticity, because these tests hinge on asymptotic properties derived from linear models, and thus lead to severe biases even in relatively large samples (Cameron & Trivedi, 2010).[21]

On this issue, strategy scholars need to address the following questions. First, if residuals are heteroscedastic, can we model such heteroscedasticity? Modeling heteroscedasticity in Tobit models is not an easy task and may be arbitrary.[22] Indeed, if the residuals are heteroscedastic (and serially correlated in panel applications), scholars cannot simply use a "robust" version of their Tobit model, because there is not a Huber-White-type estimator for Tobit models that corrects for heteroscedasticity (and serial correlation) (Greene, 2003). However, bootstrapping *SE*s may solve the issue of heteroscedasticity.[23] Generally, scholars need to consider the benefits of including time dummies (when dealing with panel data), geographic dummies (when dealing with multi-country/region samples), and industry dummies (when using cross-industry samples). These approaches alleviate problems of heteroscedasticity coming from an incorrect model specification, where some relevant regressors are omitted; because the effect of such regressors is in the error term, it may lead to heteroscedastic residuals. A more specific approach—especially when dealing with panel data—is provided by clustering. As shown, for instance, by Cameron and Miller (2015), serial correlation within clusters likely leads to a large difference between unadjusted *SE*s and clustered ones. To this extent, it is important to understand which is the most tailored *dimension* of clustering. For instance, in our application the source of heteroscedasticity and serial correlation was likely due to the industry level. In other applications the proper dimension may be, for instance, the year level, the firm level, or the geographic level. For some applications of clustered *SE*s in Tobit estimations see Eckel, Fatas, and Wilson (2010) and Jain and Thietart (2014) as well as the statistical package related to Petersen (2009) which provides two-level clustering for Tobit models.[24] We warn the reader, however, that clustering in the context of Tobit models has not received enough methodological scrutiny from a theoretical standpoint. An exception is represented by Andersen, Benn, Jørgensen, and Ravn (2013).

Second, strategy scholars need to test whether residuals are non-Normal. A preliminary step could be to graphically plot the Tobit residuals. Clearly, Tobit residuals are often non-Normal due to the censoring, but in certain applications they may be log-Normal. In these instances, a useful step is to apply a logarithmic transformation to the dependent variable (Laursen & Salter, 2006).[25] More formally, scholars may use the Stata command *tobcm* to implement a bootstrap-based conditional moment test on the null hypothesis that the residuals are Normal (for more details see Skeels & Vella, 1999 and Drukker, 2002). As shown in Table 2, this test strongly rejects Normality in our data. When implementing this test it is worth noting that: (a) it has high statistical power for samples with more than 500 observations, and (b) *tobcm* only works with left censoring at zero and no right censoring (Cameron & Trivedi, 2010; Drukker, 2002).

Finally, when heteroscedasticity or non-Normality are thought to be a serious concern for the accuracy of Tobit estimates, the censored least absolute deviations (CLAD) estimator (Powell, 1984)—a Tobit estimator which remains consistent when the residuals are asymmetrically distributed and their conditional expected median value equals zero—is a suitable choice (for more details see, for instance, Wilhelm, 2008).

4.4 | Further suggestions about model specification and estimation

In many empirical applications Tobit models are used with discrete dependent variables (see, for instance, the debate in Blundell & Smith, 1994). While Tobit models are suitable for censored dependent variables whose uncensored distribution is continuous, the use of Tobit models may be problematic when dealing with discrete dependent variables. To this extent, in the case of count variables we advise strategy scholars to use, for instance, zero-inflated models where the dependent variable follows a mixed distribution: it has a density mass at zero (following a Bernoulli distribution) and a Poisson or a Negative Binomial distribution for non-zero values (Farewell et al., 2017). Alternatively, in case of corner solutions, hurdle models for count data (Cameron & Trivedi, 2010)—in which the selection equation is estimated by means of a Probit/Logit model and the main equation is estimated by means of a (truncated at zero) count data model, such as a Poisson or a negative binomial model—represent a valid choice (see Garcia, 2013 for an empirical application using the Stata package). The main difference between zero-inflated and hurdle models is that the latter do not assume a mixed distribution for the dependent variable, but they treat the zeros and the non-zeros as coming from two distinct data generating processes (for more details see Gurmu, 1998).

Once all the questions about the nature of the dependent variable have been addressed, the researcher needs to think about whether she is interested in the observed variable or in the latent variable. This decision is key to estimate the most suitable type of marginal effects, and thus to provide the most relevant managerial or policy implications. For instance, when investigating the market for managers, the compensation for managers who are searching a new job is unobservable. Indeed, a value of zero does not mean that the manager works for zero wages. In cases like this one, the researcher is typically interested in the wage that the manager could earn if she were employed (i.e., the latent variable). However, in other situations such as those related to diversification strategies of global corporations, researchers are typically interested in the observed dependent variable. Being all global corporations allegedly "at risk of diversification," zeros are likely true zeros. As shown above, in these cases, Tobit models may constitute a valid estimation approach.

5 | CONCLUSION

A growing strand of methodological research in strategy emphasizes the importance of accurately estimating given empirical relationships in order to formulate reliable managerial implications. Contributing to this literature, we have provided a comprehensive assessment of censored data and Tobit models in strategy research. We have proposed an extensive set of guidelines and suggestions, collected in Table 3 and reported in the form of a decision tree in Figure 2, which will hopefully bring some clarity to deal with censored data in strategy research.

TABLE 3 Checklist for applying Tobit models

Theory:

- Does the dependent variable take values within certain ranges (e.g., [0, 100]) or a density mass at given points of its distribution?

- Is the dependent variable censored, or does it display a corner solution?

- Is the dependent variable only theoretically censored (e.g., theoretically bounded between 0% and 100% but without any censoring)?

- Are we correctly distinguishing between censoring, corner solution, truncation, and sample selection?

- What are the assumptions about the nature of the zeros in the data? Are true zeros representing the actual choice of the economic agents under investigation? Are zeros imputed values to missing data?

- Do the zeros and the non-zero values conceptually arise from two distinct stochastic processes rather than a common process leading to a corner solution?

- What are the theoretical or empirical processes and/or coding procedures put in place by data provider that lead to data censoring?

- What are the specific thresholds of censoring? Have such thresholds been created during the data collection process, or are they suggested by existing research?

- How relevant is censoring (i.e., proportion of censored observations) in theory given the potential distribution of our dependent variable (i.e., at what point of the data distribution is censoring likely to kick in, given our knowledge of the dependent variable)?

- Do the determinants of the binary decision (e.g., to invest or not) also explain—with the same sign—the intensity decision (e.g., how much to invest)?

- Are the unobservable factors (not included in the model) that potentially influence the binary decision uncorrelated with unobservable factors affecting the intensity decision?

Summary statistics and reporting:

- What is the sample size?

- What is the percentage of (assumed) censored observations?

- Do the observations with zeros and non-zeros systematically differ along some observable characteristics? Do *t* tests for the equality of means for all of the covariates through which zeros and non-zeros are compared reject the null hypothesis? Alternatively, does the LR test on the joint insignificance of mean covariate differences between zeros and non-zeros reject the null hypothesis?

- Are regressors correlated with both the dependent variable of the selection equation and the one of the main equation?

- Are residuals heteroscedastic? Are residuals Normal? In panel applications, are they serially correlated?
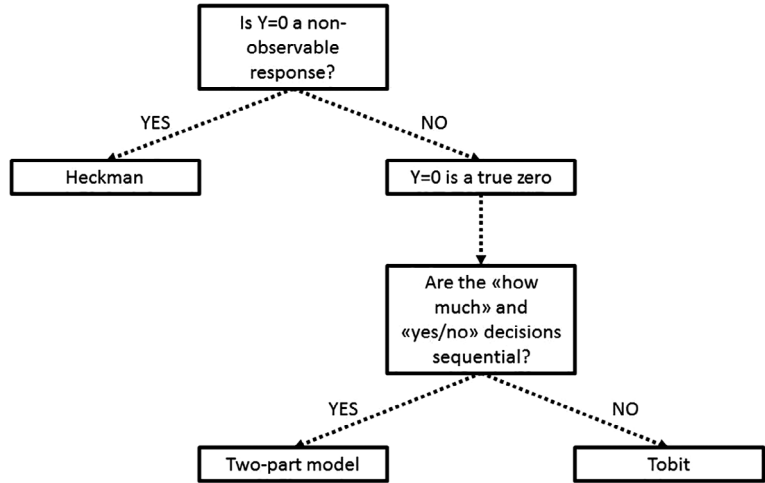
Estimation:

- Is the sample potentially affected by selection bias? Are the binary decision (e.g., to invest or not) and the intensity decision (e.g., how much to invest) sequential or simultaneous? Are the binary decision and the intensity decision independent (i.e., are two-part models preferred to Heckman models)? Does the LR test on the independence of the selection and the main equation reject the null hypothesis?

- Are we interested in estimating marginal effects on the latent variable (and thus Heckman-type models are the most suitable choice), or on the observed variable (and thus two-part models may be preferred because of less structural assumptions)?

- Is the dependent variable continuous or discrete? Is the dependent variable an uncensored proportion? In the case of count variables, are zero-inflated or hurdle models more suitable?

- If residuals are non-Normal, can we use some transformation to make them Normal?

T A B L E 3   (Continued)

- If residuals are heteroscedastic, can we model such heteroscedasticity? Does the inclusion of time dummies, geographic dummies and industry dummies solve the issue? If not, can bootstrap address it?

- Can residuals be clustered? If yes, which is the most tailored dimension of clustering?

FIGURE 2   Choice of the estimation method

ENDNOTES

[1] See Amemiya (1984), Dhrymes (1986), and Maddala (1983) for a survey.

[2] Comparing OLS with Tobit, Mudambi and Helper (1998) provide an application of how the method of moments developed in Greene (1981) can be used to adjust for the bias in OLS and thus derive results similar to maximum likelihood Tobit estimates.

[3] Scholars use interchangeably the terms "left censoring," "lower censoring," or "censoring from below" (and right censoring, upper censoring and censoring from above).

[4] In this case, there is a (known) left censoring at zero in the distribution of y^*. For the sake of simplicity, in this work we focus on left censoring but our arguments are easily generalizable to right censoring. See Carson and Sun (2007) for an extension where the censoring points are unknown.

[5] See the Appendix for details. In the Online Appendix we report the Stata commands to replicate results in Table 1 and in all tables in the Appendix.

[6] Inspecting the same journals in more recent years, we found 26 articles using Tobit in 2016, 21 in 2017, and 23 in 2018. These numbers confirm the upward trend of Tobit models in management research.

[7] Our results are largely similar if we start the analysis from 1990.

[8] We trim 1% of observations in the left and right tails of the distribution of each Compustat ratio to avoid outliers.

[9] Given the longitudinal structure of the data, we should estimate panel Tobit models; however, we prefer to keep the analysis as simple as possible and estimate pooled Tobit models. As suggested by Czarnitzki and Toole (2011), if we have the model $y_{it} = \max(0; x_{it}*\beta + c_i + \mu_{it})$, where c_i is the unobserved firm-specific effect, and assume that c_i is equal to zero, the model can be estimated as a pooled cross-sectional Tobit estimator (with clustered standard errors). Instead, if we assume that c_i is not equal to zero, the model can be estimated by means of a random-effects panel Tobit estimator. This latter hinges on the strict exogeneity assumption (i.e., the error term must be uncorrelated with the vector x_{it} across all time periods). Further, c_i must be uncorrelated with the vector x_{it}. "[D]ue to these stronger assumptions, we do not necessarily consider the panel specification as superior to the pooled cross-sectional results" (Czarnitzki & Toole, 2011, p. 152). For the sake of completeness, we verify the robustness of our results in Table 2 to the use of a random-effects Tobit regression.

[10] We use the Stata command *nehurdle* (for more details see Sánchez-Peñalver, 2019). See also the official Stata command *churdle*.

[11] There are other two types of marginal effects, which are less frequently used: the effect on (a) the conditional expected value of the dependent variable, and (b) the probability that the dependent variable is larger than the lower bound. All the four types of marginal effects can be estimated by means of the Stata command *dtobit*.

[12] The whole distribution of marginal effects in Panels A and B is available upon request. For detailed guidance on how to derive the whole distribution of marginal effects see, for instance, Wiersema and Bowen (2009).

[13] See the Section 4.4 for more details.

[14] In the context of our analysis, TNH models estimate the decision to diversify or not with a probit model, and the decision about the degree of diversification with a truncated regression (that is, the dependent variable is assumed to follow a truncated Normal distribution).

[15] In principle, researchers can run a standard Chow test on the joint insignificance of the differences across covariates between the two steps (null hypothesis) to test for the presence of two different mechanisms. However, the standard Chow test is not asymptotically valid in Tobit models (Anderson, 1987). For a consistent Chow test for Tobit models, see the procedure developed by Scott and Garen (1994).

[16] Whenever possible, we advise management scholars to employ exclusion restrictions in Heckman models.

[17] Incidental parameter problems in nonlinear panel data arise when estimators fail to converge to consistent estimates as the number of observations becomes large. Assuming N firms and T time periods, in linear models the N firm fixed effects can be differenced out (by means of, for instance, within-group estimation), and thus are not estimated. By contrast, in nonlinear models the use of firm fixed effects typically requires the additional estimation of N-1 coefficients and their correlation with the regressors in the model specification. The inclusion of these additional regressors may distort the shape of the likelihood function, and its maximization may incur in unreliable numerical solutions. For details see Greene (2004) and Lancaster (2000).

[18] Formally, scholars need to frame the problem in terms of a double-hurdle approach, that is, subjects "must pass two hurdles before being observed with a positive level of consumption" (Madden, 2008, p. 301), and these two hurdles are a "yes/no" decision about doing a certain activity and, in the case of a "yes," a decision on "how much" effort to dedicate to such an activity. As Madden (2008) argues, if the residuals of the equations modeling the two hurdles are independent, the double-hurdle model "collapses" to the Cragg model (Cragg, 1971).

[19] To test whether Heckman and two-part models display the same explanatory power, researcher may use the Vuong's (1989) LR test for non-nested estimators (for an application see Tomlin, 2000).

[20] Namely, unconditional marginal effects can be calculated by combining the estimated average probability from the probit model with the OLS coefficients.

[21] Cameron and Trivedi (2010) explain in detail why standard tests are biased, and report a step-by-step procedure through which researchers can test Normality and homoscedasticity in Tobit models.

[22] Heteroscedasticity can also be modeled by means of the Stata command *intreg*.

[23] It is worth noting that the effectiveness of bootstrap may depend on the sample size. As argued by Guan (2003): "While the nonparametric bootstrap method does not rely upon strong assumptions regarding the distribution of the statistic, a key assumption of bootstrapping is the similarity between the characteristics of the sample and of the population. When the sample is of size 500 (100 independent clusters), the assumption of similarity may not be reasonable. […] In summary, the number of repetitions and sample size both play important roles in the bootstrap method" (p. 80).

[24] Computing cluster-robust standard errors can be problematic in the case of a low number of clusters.

[25] In the case of elliptically contoured distributions of residuals, please see the methodology in Barros, Galea, Leiva, and Santos-Neto (2018).

[26] While we focus on cross-section analysis, our findings are also useful for panel analysis; see Czarnitzki and Toole (2011) for a discussion.

[27] It is worth noting that in our setting we do not need to set the values of coefficients of our regressors to mimic *small* or *large* effects—such as in the literature on effect sizes (Cohen, 1992; Cohen, Cohen, West, & Aiken, 2013).

Indeed, by manipulating the properties of the dependent variable (presence of selection bias) and residuals (heteroscedasticity; distribution function), we generate the *true* effect in our datasets and use it as reference point to calculate the bias associated with estimates.

[28] ε_1 and ε_2 are assumed to be bivariate Normal, with mean zero and covariance matrix equal to $\begin{bmatrix} \sigma_{\varepsilon_1} & \rho \\ \rho & \sigma_{\varepsilon_2} \end{bmatrix}$, where $y_2 = y_2^*$ if $y_1 = 1$ (y_2 is observed only when y_1 is not zero). Equation (A3) is estimated by means of a Probit estimator—y_1 is a dummy variable—while Equation (A2) is estimated by means of an OLS regression with the inclusion of the inverse Mills' ratio term (also called Heckman's lambda).

[29] In the estimation of the Heckman selection model, we do not employ any exclusion restriction. Note that not employing any exclusion restriction means that the vector of regressors explaining the dependent variable in Equation (A2) is the same vector explaining the dependent variable in Equation (A3).

[30] As Lumley et al. (2002, p. 152) show, for sufficiently large samples OLS "rely on the Central Limit Theorem, which states that the average of a large number of independent random variables is approximately Normally distributed around the true population mean. It is this Normal distribution of an average that underlies the validity of the t test and linear regression." This is important when the mean is the primary goal of estimation.

[31] Specifically, focusing on the latent variable we find that the bias of Tobit models exists even in a dataset of 10 million observations, ranging between 5.44% ($\delta E(y^*)/\delta x_1$) and 5.5% ($\delta E(y^*)/\delta x_2$); it becomes extremely large (between 10.28% ($\delta E(y^*)/\delta x_1$) and 18.83% ($\delta E(y^*)/\delta x_2$)) when the number of observations is small (Panel B). Focusing on the observed variable, we find the same pattern but smoother: the bias of Tobit models in the dataset of 10 million observations (Panel A) ranges between 2.11% ($\delta E(y)/\delta x_2$) and 2.28% ($\delta E(y)/\delta x_1$), while in the small dataset (Panel B) the bias exceeds 13%.

REFERENCES

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?* NBER Working Paper No. w24003.

Aitchison, J. (1955). On the distribution of a positive random variable having a discrete probability mass at the origin. *Journal of the American Statistical Association*, *50*(271), 901–908.

Alan, S., Honoré, B. E., Hu, L., & Leth-Petersen, S. (2014). Estimation of panel data regression models with two-sided censoring or truncation. *Journal of Econometric Methods*, *3*(1), 1–20.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics*, *24*, 3–61.

Andersen, A., Benn, C. S., Jørgensen, M. J., & Ravn, H. (2013). Censored correlated cytokine concentrations: Multivariate Tobit regression using clustered variance estimation. *Statistics in Medicine*, *32*(16), 2859–2874.

Anderson, G. J. (1987). Prediction tests in limited dependent variable models. *Journal of Econometrics*, *34*(1–2), 253–261.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics*, *19*(1), 2–28.

Arabmazar, A., & Schmidt, P. (1981). Further evidence on the robustness of the Tobit estimator to heteroskedasticity. *Journal of Econometrics*, *17*(2), 253–258.

Arabmazar, A., & Schmidt, P. (1982). An investigation of the robustness of the Tobit estimator to non-Normality. *Econometrica*, *50*(4), 1055–1063.

Barros, M., Galea, M., Leiva, V., & Santos-Neto, M. (2018). Generalized Tobit models: Diagnostics and application in econometrics. *Journal of Applied Statistics*, *45*(1), 145–167.

Baum, C. F. (2008). Stata tip 63: Modeling proportions. *Stata Journal*, *8*(2), 299–303.

Becker, W., & Dietz, J. (2004). R&D cooperation and innovation activities of firms – Evidence from the German manufacturing industry. *Research Policy*, *33*(2), 209–223.

Bernard, A., Jensen, J. B., & Schott, P. (2006). Survival of the best fit: Low wage competition and the (uneven) growth of US manufacturing plants. *Journal of International Economics*, *68*, 219–237.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, *119*(1), 249–275.

Blevins, D. P., Tsang, E. W., & Spain, S. M. (2015). Count-based research in management: Suggestions for improvement. *Organizational Research Methods*, *18*(1), 47–69.

Blundell, R., & Smith, R. J. (1994). Coherency and estimation in simultaneous models with censored or qualitative dependent variables. *Journal of Econometrics*, *64*(1–2), 355–373.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*(430), 443–450.

Bowen, H. P., & Wiersema, M. F. (2004). Modeling limited dependent variables: Methods and guidelines for researchers in strategic management. In *Research methodology in strategy and management* (Vol. 1). Bingley, UK: Emerald Group Publishing.

Bowen, H. P., & Wiersema, M. F. (2005). Foreign-based competition and corporate diversification strategy. *Strategic Management Journal*, *26*(12), 1153–1171.

Brammer, S., & Millington, A. (2008). Does it pay to be different? An analysis of the relationship between corporate social and financial performance. *Strategic Management Journal*, *29*(12), 1325–1343.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using Stata*. College Station, TX: Stata Press.

Carson, R., & Sun, Y. (2007). The Tobit model with a non-zero threshold. *The Econometrics Journal*, *10*(3), 488–502.

Cassiman, B., & Veugelers, R. (2006). In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science*, *52*(1), 68–82.

Certo, S. T., Busenbark, J., Woo, H. S., & Semadeni, M. (2016). Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal*, *37*(13), 2639–2657.

Clougherty, J. A., Duso, T., & Muck, J. (2016). Correcting for self-selection based endogeneity in management research: Review, recommendations and simulations. *Organizational Research Methods*, *19*(2), 286–347.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Routledge.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, *39*(5), 829–844.

Czarnitzki, D., & Toole, A. A. (2011). Patent protection, market uncertainty, and R&D investment. *Review of Economics and Statistics*, *93*(1), 147–159.

Delios, A., & Beamish, P. W. (1999). Geographic scope, product diversification, and the corporate performance of Japanese firms. *Strategic Management Journal*, *20*(8), 711–727.

Devers, C. E., McNamara, G., Haleblian, J., & Yoder, M. E. (2013). Do they walk the talk? Gauging acquiring CEO and director confidence in the value creation potential of announced acquisitions. *Academy of Management Journal*, *56*(6), 1679–1702.

Dhrymes, P. J. (1986). Limited dependent variables. In Griliches & Michael D. Intriligator (Eds.), *Handbook of econometrics* (Vol. 3, pp. 1567–1631). Holland: Elsevier.

Dow, W. H., & Norton, E. C. (2003). Choosing between and interpreting the Heckit and two-part models for corner solutions. *Health Services and Outcomes Research Methodology*, *4*(1), 5–18.

Drukker, D. M. (2002). Bootstrapping a conditional moments test for normality after Tobit estimation. *Stata Journal*, *2*(2), 125–139.

Duan, N., Manning, W. G., Morris, C. N., & Newhouse, J. P. (1983). A comparison of alternative models for the demand for medical care. *Journal of Business & Economic Statistics*, *1*(2), 115–126.

Eckel, C. C., Fatas, E., & Wilson, R. (2010). Cooperation and status in organizations. *Journal of Public Economic Theory*, *12*(4), 737–762.

Farewell, V. T., Long, D. L., Tom, B. D. M., Yiu, S., & Su, L. (2017). Two-part and related regression models for longitudinal data. *Annual Review of Statistics and its Application*, *4*, 283–315.

García, B. (2013). Implementation of a double-hurdle model. *The Stata Journal*, *13*(4), 776–794.

Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric Reviews*, *23*(2), 125–147.

Greene, W. H. (1981). On the asymptotic bias of the ordinary least square estimator of the Tobit model. *Econometrica*, *49*, 505–514.

Greene, W. H. (2003). *Econometric analysis*. Upper Saddle River, NJ: Prentice Hall.

Guan, W. (2003). From the help desk: Bootstrapped standard errors. *Stata Journal*, *3*(1), 71–80.

Gurmu, S. (1998). Generalized hurdle count data regression models. *Economics Letters*, *58*(3), 263–268.

Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, *1*(1), 51–78.

Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, *28*(4), 331–343.

Honoré, B. E. (1992). Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica*, *60*(3), 533–565.

Humphreys, B. R. (2013). Dealing with zeros in economic data. Working Paper.

Jain, A., & Thietart, R. A. (2014). Capabilities as shift parameters for the outsourcing decision. *Strategic Management Journal*, *35*(12), 1881–1890.

Jones, A. M. (2000). Health econometrics. In Anthony J. Culyer & Joseph P. Newhouse (Eds.), *Handbook of health economics* (Vol. 1, pp. 265–344). Holland: Elsevier.

Kennedy, P. (2008). *A guide to modern econometrics*. Oxford, England: Blackwell Publishing.

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, *17*, 441–458.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, *95*(2), 391–413.

Laursen, K., & Salter, A. (2006). Open for innovation: The role of openness in explaining innovation performance among UKmanufacturing firms. *Strategic Management Journal*, *27*(2), 131–150.

Leung, S. F., & Yu, S. (1996). On the choice between sample selection and two-part models. *Journal of Econometrics*, *72*(1–2), 197–229.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*(1), 151–169.

Luo, Y., & Bu, J. (2018). When are emerging market multinationals more risk-taking? *Global Strategy Journal*, *8*(4), 635–664.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge, England: Cambridge University Press.

Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics*, *27*(2), 300–307.

Molina-Azorin, J. F. (2012). Mixed methods research in strategic management impact and applications. *Organizational Research Methods*, *15*(1), 33–56.

Mudambi, R., & Helper, S. (1998). The "close but adversarial" model of supplier relations in the US auto industry. *Strategic Management Journal*, *19*, 775–792.

Papke, L. E., & Wooldridge, J. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*, 619–632.

Papke, L. E., & Wooldridge, J. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*, 121–133.

Pepper, J. V. (2002). Robust inferences from random clustered samples: An application using data from the panel study of income dynamics. *Economics Letters*, *75*(3), 341–345.

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

Peterson, M., Arregle, J. L., & Martin, X. (2012). Multilevel models in international business research. *Journal of International Business Studies*, *43*, 451–457.

Powell, J. L. (1984). Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, *25*(3), 303–325.

Ragozzino, R., & Reuer, J. J. (2011). Geographic distance and corporate acquisitions: Signals from IPO firms. *Strategic Management Journal*, *32*(8), 876–894.

Reuer, J. J., & Leiblein, M. J. (2000). Downside risk implications of multinationality and international joint ventures. *Academy of Management Journal*, *43*(2), 203–214.

Sánchez-Peñalver, A. (2019). Estimation methods in the presence of corner solutions. *Stata Journal*, *19*(1), 87–111.

Scott, F., & Garen, J. (1994). Probability of purchase, amount of purchase, and the demographic incidence of the lottery tax. *Journal of Public Economics*, *54*(1), 121–143.

Semadeni, M., Withers, M. C., & Certo, S. T. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, *35*(7), 1070–1079.

Shook, C. L., Ketchen, D. J., Cycyota, C. S., & Crockett, D. (2003). Data analytic trends and training in strategic management. *Strategic Management Journal*, *24*(12), 1231–1237.

Shook, C. L., Ketchen, D. J., Hult, G. T. M., & Kacmar, K. M. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal*, *25*(4), 397–404.

Silva, J. M. S., Tenreyro, S., & Windmeijer, F. (2015). Testing competing models for non-negative data with many zeros. *Journal of Econometric Methods*, *4*(1), 29–46.

Skeels, C. L., & Vella, F. (1999). A Monte Carlo investigation of the sampling behavior of conditional moment tests in Tobit and Probit models. *Journal of Econometrics*, *92*(2), 275–294.

Sorenson, O., McEvily, S., Ren, C. R., & Roy, R. (2006). Niche width revisited: Organizational scope, behavior and performance. *Strategic Management Journal*, *27*(10), 915–936.

Tan, B. R., & Chintakananda, A. (2016). The effects of home country political and legal institutions on firms' geographic diversification performance. *Global Strategy Journal*, *6*(2), 105–123.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, *26*(1), 24–36.

Tomlin, K. M. (2000). The effects of model specification on foreign direct investment models: An application of count data models. *Southern Economic Journal*, *67*(2), 460–468.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, *57*(2), 307–333.

Wiersema, M. F., & Bowen, H. P. (2009). The use of limited dependent variable techniques in strategy research: Issues and methods. *Strategic Management Journal*, *30*(6), 679–692.

Wilhelm, M. O. (2008). Practical considerations for choosing between Tobit and SCLS or CLAD estimators for censored regression models with an application to charitable giving. *Oxford Bulletin of Economics and Statistics*, *70*(4), 559–582.

Wolfolds, S. E., & Siegel, J. (2019). Misaccounting for endogeneity: The peril of relying on the Heckman two-step method without a valid instrument. *Strategic Management Journal*, *40*(3), 432–462.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wulff, J. N., & Villadsen, A. R. (2019). Keeping it within bounds: Regression analysis of proportions in international business. *Journal of International Business Studies*, in press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

A. APPENDIX

Simulations are a common approach to understand how an estimation method performs against a given baseline. As Certo et al. (2016) discuss (see also Kennedy, 2008; Semadeni et al., 2014), simulations involve two main steps: (a) generating datasets with known properties, and (b) analyzing such

datasets with different models. Following these indications, we use a Monte Carlo method to build three datasets of different size: 10 million, 100 and 1,000 independent observations, respectively. The median sample size in our review of management articles is 894 observations, which is close to the third dataset used in our simulation. Then, we manipulate the values of y^* and y, the covariates (x_1 and x_2) and the residuals (ε) to reproduce different conditions about selection bias, non-Normality and heteroscedasticity of residuals.[26] As for x_1 and ε, we build them as Normally distributed random variables with means and standard deviations equal to zero and one respectively. As for x_2, we build it as a uniformly distributed random variable in the interval [0,1]. The functional form of y^* is the following:

$$y^* = 1 + x_1 + x_2 + \varepsilon \tag{A1}$$

For the sake of simplicity, we choose a simple linear specification where y^* is the sum of a constant (equal to unity), x_1, x_2 and ε.[27]

Descriptive statistics are reported in Table A1. As shown, whichever the sample size, y^* and y have the same median value. However, the mean value of y^* is lower than that of y because of their different range values: while y^* can assume negative values, y cannot. This different range also explains the lower standard deviation in the distribution of y as compared to y^*. As for x_1 and ε, in the dataset of 10 million observations such variables are very close to the hypothesized standard Normal distribution. Also the mean and standard deviation of x_2 are close to the hypothesized ones in the case of the large dataset.

A.1. Censoring versus selection bias

An empirical complication with Tobit models is an apparent interchangeability between censoring and selection bias. Some previous studies implicitly assumed that Tobit could solve self-selection bias because it allows to investigate both a given *binary* decision (e.g., diversify or not the business), and the related decision about the *intensity* (e.g., how much to diversify). However, Tobit models—when used to address sample selection (rather than for censoring or corner solutions)—impose that the explanatory variables explaining the decision to diversify must also explain—*with the same sign*—the intensity of diversification. Here, we show that using Tobit to address sample selection bias is an inferior approach as compared to an Heckman selection model. We have a system of two equations where Equation (A2) is the main equation and Equation (A3) is the selection equation[28]:

$$y_2^* = X'\beta + \varepsilon_2 \tag{A2}$$

$$y_1 = Z'\gamma + \varepsilon_1 \tag{A3}$$

In other words, the dependent variable in the main equation (e.g., R&D intensity) is observed if and only if the dependent dichotomous variable in the selection equation (e.g., dummy variable indicating whether the firm engages in R&D activities) equals one. A Tobit estimation imposes a rigid constraint in that the regressors that drive the choice to, for instance, invest in R&D or not must also explain the magnitude of such an R&D investment. If the two decisions—whether investing in R&D and, if so, how much—are explained by different factors, then a standard Tobit model could not be appropriate.

TABLE A1   Descriptive statistics

Panel A. Sample size = 10,000,000					
	Mean	**Median**	**SD**	**Min**	**Max**
y^*	1.4994	1.4997	1.4438	−6.2035	9.4988
y	1.6112	1.4997	1.2621	0	9.4988
x_1	−0.0001	0.0001	1.0003	−5.0653	5.3122
x_2	0.4999	0.4998	0.2886	0.0000	1.0000
ε	−0.0003	−0.0008	1.0001	−5.3221	5.1942
Panel B. Sample size = 100					
	Mean	**Median**	**SD**	**Min**	**Max**
y^*	1.8264	1.9495	1.7596	−2.7210	6.0156
y	1.9707	1.9495	1.5153	0	6.0156
x_1	0.1142	0.0082	1.1444	−3.1819	2.6153
x_2	0.5602	0.5800	0.2926	0.0071	0.9965
ε	0.1520	0.1245	1.0818	−2.1165	2.4191
Panel C. sample size = 1,000					
	Mean	**Median**	**SD**	**Min**	**Max**
y^*	1.5341	1.5804	1.4344	−2.7051	6.3536
y	1.6452	1.5804	1.2474	0	6.3536
x_1	0.0652	0.0380	1.0105	−3.1819	2.6968
x_2	0.4859	0.4823	0.2884	0.0003	0.9993
ε	−0.0169	0.0228	0.9986	−3.5409	3.4484

In Table A2, we compare Heckman selection and Tobit estimates.[29] As above, the comparison of the marginal effects in Column I (II) and Column III (IV) hinges on the fact that the Heckman selection regression in Column I (II) estimates *by construction* the true effect in the simulated data. The difference between the two marginal effects represents the distortion in Tobit estimates when the above assumption does not hold. In Panels A, B, and C, we show the marginal effects based on the datasets of 10 million, 100 and 1,000 observations, respectively. As shown, when the sample is large (Panel A) there are small differences between Tobit and Heckman selection estimates. When the sample is small (Panel B), such differences become larger.

A.2. Non-normality and heteroscedasticity of residuals

Tobit models rely on the assumption that residuals are homoscedastic and Normal (e.g., Arabmazar & Schmidt, 1981, 1982). Yet, as shown by Brammer and Millington (2008) residuals of Tobit estimations are often likely to violate this assumption: for instance, Normality and homoscedasticity are almost certainly violated due to a large number of zeros. While some authors (e.g., Laursen & Salter, 2006) explicitly *assume* a distribution for the residuals in their Tobit model specification, a large number of reviewed studies does not seem to deal at all with the distribution of Tobit residuals.

TABLE A2   Selection bias

Panel A. Sample size = 10,000,000; censored Obs. = 15%

	$\delta E(y^*)/\delta X$ Heckman	$\delta E(y)/\delta X$ Heckman	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	0.9990	0.9334	1.0005	0.9335
	(0.0007)	(0.0006)	(0.0003)	(0.0003)
x_2	0.9984	0.9328	0.9998	0.9329
	(0.0013)	(0.0012)	(0.0011)	(0.0011)
Pseudo R^2			.2099	

Panel B. Sample size = 100; censored Obs. = 15%

	$\delta E(y^*)/\delta X$ Heckman	$\delta E(y)/\delta X$ Heckman	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	1.0734	1.0526	1.1213	1.0707
	(0.1654)	(0.1505)	(0.1061)	(0.1013)
x_2	1.1378	1.1149	1.1812	1.1278
	(0.4131)	(0.3952)	(0.3895)	(0.3719)
Pseudo R^2			.2478	

Panel C. Sample size = 1,000; censored Obs. = 14.7%

	$\delta E(y^*)/\delta X$ Heckman	$\delta E(y)/\delta X$ Heckman	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	1.0677	0.9171	0.9855	0.9264
	(0.0967)	(0.0800)	(0.0334)	(0.0314)
x_2	0.9322	0.7968	0.8624	0.8107
	(0.1754)	(0.1592)	(0.1112)	(0.1045)
Pseudo R^2			.2066	

Note: Regressions are estimated with an intercept term. *SE*s are reported in parentheses.

While OLS regressions may be valid for non-Normally distributed outcomes,[30] Tobit models are not. To highlight this problem, we assume that OLS estimates are consistent and we simulate ε as a Poisson distributed random variable, and compare OLS and Tobit estimates (in Table A3). In Panels A, B, and C, we show the marginal effects based on the datasets of 10 million, 100 and 1,000 independent observations, respectively. As shown, when residuals are non-Normal (and the dependent variable has around half of censored observations) the bias of Tobit models can be quite relevant especially when working with small samples.[31]

While in OLS regressions heteroscedastic errors lead to Normally distributed estimated coefficients (Lumley et al., 2002), in Tobit models heteroscedasticity must be explicitly modeled to avoid an inconsistent estimation of marginal effects. To this extent, we model ε as:

$$\eta \mid X \sim N\left(0; \sigma^2 \exp(kw)\right) \tag{A4}$$

TABLE A3 Non-normality of residuals

	$\delta E(y^*)/\delta X$ OLS	$\delta E(y)/\delta X$ OLS	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
Panel A. Sample size = 10,000,000; censored observations = 52.2%				
x_1	1.0000	0.4777	1.0544	0.4668
	(0.0002)	(0.0002)	(0.0005)	(0.0002)
x_2	1.0009	0.4776	1.0559	0.4675
	(0.0008)	(0.0007)	(0.0013)	(0.0006)
R^2	.6501	.3811		
Pseudo R^2			.2744	
Panel B. Sample size = 100; censored observations = 49%				
x_1	0.9039	0.4419	0.8110	0.4375
	(0.0589)	(0.0482)	(0.0881)	(0.0475)
x_2	1.5011	0.8511	1.7837	0.9621
	(0.2305)	(0.1885)	(0.3260)	(0.1758)
R^2	.7569	.5413		
Pseudo R^2			.3812	
Panel C. Sample size = 1,000; censored observations = 48.7%				
x_1	1.0299	0.5138	1.1188	0.5326
	(0.0255)	(0.0218)	(0.0469)	(0.0223)
x_2	0.9539	0.4994	1.0449	0.4974
	(0.0894)	(0.0764)	(0.1303)	(0.0620)
R^2	.6316	.3706		
Pseudo R^2			.2677	

Note: Regressions are estimated with an intercept term. *SE*s are reported in parentheses.

where k is a constant (set equal to 0.5) and w is a randomly distributed dummy variable across observations. In Table A4, we compare OLS and Tobit estimates. Focusing on the marginal effects calculated on the observed variable, and assuming that OLS estimates are consistent, in Panel A we show that the bias associated with Tobit estimates is quite large even in large datasets (just below 7%). Yet, the bias can become more relevant (up to 10.3%) when we work with the smaller samples.

TABLE A4 Heteroscedasticity of residuals

Panel A. Sample size = 10,000,000; censored observations = 18.5%			
	$\delta E(y^*)/\delta X$ OLS	$\delta E(y)/\delta X$ OLS	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	0.9997	0.8147	1.0022	0.8708
	(0.0004)	(0.0004)	(0.0005)	(0.0004)
x_2	0.9988	0.8159	1.0018	0.8704
	(0.0015)	(0.0013)	(0.0015)	(0.0013)
R^2	.3683	.3484		
Pseudo R^2			.1261	

Panel B. Sample size = 100; censored observations = 15%			
	$\delta E(y^*)/\delta X$ OLS	$\delta E(y)/\delta X$ OLS	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	1.0819	0.8606	1.0313	0.9491
	(0.1191)	(0.0997)	(0.1188)	(0.1093)
x_2	1.0473	0.9355	1.0875	1.0008
	(0.4660)	(0.3899)	(0.4411)	(0.4059)
R^2	.4891	.4689		
Pseudo R^2			.1831	

Panel C. Sample size = 1,000; censored observations = 16.9%			
	$\delta E(y^*)/\delta X$ OLS	$\delta E(y)/\delta X$ OLS	$\delta E(y^*)/\delta X$ Tobit	$\delta E(y)/\delta X$ Tobit
x_1	0.9589	0.7840	0.9534	0.8448
	(0.0419)	(0.0362)	(0.0442)	(0.0392)
x_2	1.1806	0.9749	1.1278	0.9994
	(0.1468)	(0.1270)	(0.1483)	(0.1314)
R^2	.3656	.3403		
Pseudo R^2			.1225	

Note: Regressions are estimated with an intercept term. *SE*s are reported in parentheses.