




Addressing Endogeneity Without Instrumental Variables: An Evaluation of the Gaussian Copula Approach for Management Research

Christine Eckert

University of Technology Sydney

Jan Hohberger 

University Ramon Llull

The availability and quality of instrumental variables (IV) are frequent concerns in empirical management research when trying to overcome endogeneity problems. For endogeneity that does not arise from sample selection, management scholars have recently started to apply the Gaussian Copula (GC) approach as an alternative to IV regression. Although the GC approach has various promising features, its limitations and usefulness in a management context are still not fully understood. We discuss the GC approach as a flexible, instrument-free approach to correct for endogeneity and examine its suitability for applied management research. We use simulations to explore the limitations and practical usefulness of the GC approach relative to ordinary least squares (OLS), IV regression, and a Higher Moments (HM) estimator by simulating the impact of different degrees of violation of the key underlying assumptions of the GC approach. We show that the GC approach can recover the true parameters remarkably well if all of its assumptions are met but that its absolute and relative performance in terms of parameter recovery and estimation precision can deteriorate quickly if these assumptions are violated. This is of particular concern as some of these assumptions are not testable and violations of them are likely in many empirical management contexts. Based on our results, we provide a series of recommendations and practical guidelines for scholars who consider using the GC approach when dealing with endogeneity.

Acknowledgments: We would like to thank Peter Ebbes, Dominik Pappies, Ken Frank, John Busenbark, and Daniel Blaseg for sharing their knowledge with us. We also thank three anonymous reviewers and the editor Donald Bergh; all of their invaluable suggestions and feedback undoubtedly shaped and contributed to the quality of this paper. We are also very grateful for the excellent research assistance by Charles Howell and Yu Yang.

Supplemental material for this article is available with the manuscript on the JOM website.

Corresponding Author: Christine Eckert, UTS Business School, University of Technology Sydney, 14-28 Ultimo Rd, Ultimo NSW 2007, Australia.

Email: Christine.Eckert@uts.edu.au

Keywords: *endogeneity; Gaussian Copula; instrumental variable; simulation; research methods; Higher Moments*

Introduction

The correct estimation of size and directionality of relationships is central to theory testing and development; it also lies at the heart of any meaningful guidance to practitioners (Hamilton & Nickerson, 2003). Due to the complex nature of management data, however, it is often difficult to establish causality and estimate unbiased effects in empirical management research (Shaver, 2020). Researchers often face endogeneity, a situation in which the independent variable is correlated with the error term. Endogeneity can occur due to sample selection, measurement error, omitted variables, or simultaneous causality (Hill, Johnson, Greco, O'Boyle, & Walter, 2021) and can result in biased estimates and incorrect inferences.

Due to the problematic impact endogeneity may have on managerial inference as well as concerns about its existence in a wide range of settings, dealing with endogeneity has become a focal point of methodological discussions (e.g., Certo, Busenbark, Woo, & Semadeni, 2016; Hamilton & Nickerson, 2003; Hill et al., 2021). One of the most frequently prescribed and applied ways of addressing endogeneity is the instrumental variables (IV) approach (see Hill et al., 2021). It has the advantage of being applicable for a variety of sources of endogeneity, and its underlying statistical theory is well developed. However, researchers frequently encounter problems in its application as suitable IV are hard to find (Rossi, 2014). Instruments must be correlated with the endogenous regressor through a relationship that the researcher can explain (strength) and uncorrelated with the error term (exogeneity; Semadeni, Withers, & Certo, 2014). Not only is the testing of these assumptions difficult, but estimation with unsuitable instruments can lead to biased results, which may be inferior to those obtained by uncorrected ordinary least squares (OLS) regression (Murray, 2006; Semadeni et al., 2014).

To address these challenges, Park and Gupta (2012) introduced the Gaussian Copula (GC) approach as an instrument-free method of dealing with endogeneity. They discussed the GC approach in the context of omitted variables and simultaneous causality, but it has since also been applied to endogeneity arising from measurement error. Like other instrument-free approaches, such as the Higher Moments (HM) approach (Ebbes, Wedel, & Böckenholt, 2009; Lewbel, 1997), the Identification through Heteroscedasticity estimator (Lewbel, 2012; Rigobon, 2003), and the Latent Instrumental Variables method (Ebbes, Wedel, Böckenholt, & Steerneman, 2005), the GC approach does not need observed instruments to correct for endogeneity but relies on distributional assumptions to identify the impact of the variable of interest.

While the other instrument-free methods have seen only a very limited uptake in management research (e.g., Andries & Hünermund, 2020; Hasan, Taylor, & Richardson, 2021)—most likely due to their inherent complexity—the number of applications of the GC approach has recently been increasing (e.g., Becerra & Markarian, 2021; Bhattacharya, Good, Sardashti, & Pelozo, 2020; Haschka & Herwartz, 2020; Reck, Fliaster, & Kolloch, 2021). This is not surprising, since the GC approach has multiple features that may make it appealing for management research: It offers researchers an alternative when suitable instruments are unavailable;

it can be used as a robustness test when instruments are weak or their exogeneity is unknown; it can be applied to nonlinear models (e.g., Random Coefficient Logit with structural error from the normal distribution); and it can handle multiple endogenous regressors as well as interactions (Papies, Ebbes, & Van Heerde, 2016). These situations are common in the management context, but they usually provide additional challenges for the application of the classical IV approach.

The GC approach is, however, not without limitations as its ability to recover true parameters hinges on four conditions. First, the method requires the endogenous regressor to be continuous or have a certain number of discrete outcomes, implying that—apart from endogeneity due to sample selection—it also cannot handle endogeneity due to self-selection into a few discrete outcomes. Second, the endogenous regressor needs to be sufficiently nonnormal to enable the identification of the model. Third, the error term is assumed to be normally distributed. And fourth, it requires that the dependency between the distribution of the endogenous regressor and the error term can be described by a GC. Although recent management studies using the GC approach frequently test for nonnormality of the regressor (e.g., Becerra & Markarian, 2021; Bhattacharya et al., 2020), only very few studies discuss the assumption regarding the normality of the error (e.g., Becerra & Markarian, 2021) or regarding the dependency structure between the endogenous regressor and the error term (e.g., Bhattacharya et al., 2020). A clear understanding of these assumptions and their implied limitations is, however, important: Without such understanding, the GC approach may share the same fate as other methods of dealing with endogeneity that were found to be often applied incorrectly, resulting in the cure being worse than the illness (see Semadeni et al., 2014).

Park and Gupta (2012) and subsequent studies have started to discuss and explore the scope and limitations of the GC approach (e.g., Becker, Proksch, & Ringle, 2021), but the picture is still rather incomplete. The aim of this paper is thus to explore the boundaries of the GC approach in the context of management research and to provide guidance for its potential application therein. We use simulations to provide a more detailed examination of the assumptions of the GC approach and evaluate its relative performance in terms of estimation bias and precision compared to OLS, IV regression, and an alternative instrument-free method, namely the HM approach developed for omitted variable endogeneity and discussed in Ebbes et al. (2009). More specifically, we focus on variations in the distribution of the error term and the endogenous regressor to reflect scenarios frequently encountered in management research. We also explore violations of the assumed dependency structure between regressor and error term and examine the GC's relative performance in this context.

We show that the GC's performance strongly relates to how well its assumptions are met. Under ideal conditions, the GC approach provides estimates that are almost as close to the true values as those obtained with exogenous instruments of medium strength, and its power surpasses the power of IV regressions with exogenous and medium strength instruments. If all its assumptions are met, the GC approach also performs significantly better than the HM approach, which provides biased estimates in cases of medium endogeneity.

For violations of the distributional assumptions regarding the error term or the regressor, we find that the GC approach provides estimates that are usually more biased than those obtained with exogenous instruments of low strength but outperform those obtained with OLS or

endogenous instruments of high strength. In general, we find that the GC method provides estimates that are less than 1 standard error away from the true parameters, yet because these standard errors are rather large, the method's ability to detect significant relationships is impaired. In all these simulations, and similar to other instrument-free techniques (Lewbel, 2012), the GC approach is likely to give noisier, less reliable estimates than IV regression with exogenous instruments of medium strength. Compared to the HM approach, the GC approach yields less biased but also noisier estimates.

We further show that violations of the dependency structure can have a significant impact on the GC approach's ability to recover the true parameters. Using an agnostic nonparametric data-generating process, our results reveal that the estimated parameter is more than 2 standard errors away from the true parameters and more biased than OLS estimates. Although the GC approach shares this feature with the HM approach, the GC approach's bias is stronger and aggravated at higher levels of endogeneity.

Overall, the GC approach may provide a practical alternative when addressing endogeneity, particularly in situations where observed instruments are weak or of questionable exogeneity. We do, however, caution researchers to carefully consider the suitability of its underlying assumptions and particularly the assumption regarding the dependency structure between the error term and the regressor. To guide researchers in this endeavor, we discuss how this dependency structure may look depending on the underlying cause of endogeneity.

Our paper contributes in three distinct ways to the ongoing debate on endogeneity in management research (e.g., Busenbark, Yoon, Gamache, & Withers, 2022; Certo et al., 2016; Hill et al., 2021; Semadeni et al., 2014). First, we advance the discussion concerning the usefulness and applicability of the GC approach. We review previous applications of the approach and the criticisms that have emerged and discuss situational characteristics that can make the GC approach less useful or even not applicable at all. Second, we provide practical advice for the application of the GC approach and link it to previous discussions of endogeneity in management research. We use nontechnical language and focus on practical concerns of applied management scholars. Third, to our knowledge, this is the first paper that benchmarks the GC approach against standard OLS and IV regressions as well as the HM approach as an alternative instrument-free method. We show that the GC approach should be viewed as a complementary rather than as a standalone tool for dealing with endogeneity. By providing two real data examples, a step-by-step guide, and relevant STATA code, we facilitate the implementation of the GC approach so that other researchers can add it to their toolbox.

The remainder of the paper is organized in three sections. After discussion and definition of the endogeneity problem, we provide an intuitive and relatively nontechnical introduction to the GC approach. Then we draw on simulations to examine the performance of the estimator based on various distributions of the endogenous regressor and error term as well as under a different data-generating mechanism. We compare the GC approach to OLS and IV regression, as well as to the HM approach as another instrument-free estimation method, to examine its relative performance. This is followed by an application of the GC approach to two datasets from the areas of corporate social responsibility (CSR) and innovation context, respectively. We conclude with a summary of our results and recommendations for the application of the GC approach in the management context.

Methodological Background

Introduction to GC Approach

Endogeneity is defined as a situation in which the independent variable x in a linear regression

$$y = \alpha + \beta x + \varepsilon \quad (1)$$

is correlated with the unobserved error term ε . Endogeneity can result in biased estimates of $\theta = (\alpha, \beta)$ —that is, the expected value of the estimator is different from the true population parameter—and thus can yield misleading interpretations of the data (Bascle, 2008).

The starting point for the GC approach is the observation that a linear regression can be estimated using either OLS or maximum likelihood estimation and that both methods are equivalent provided that the assumptions of the linear regression model hold. Whereas in OLS parameters $\theta = (\alpha, \beta)$ are obtained using $\theta = \text{inv}(x'x)x'y$, the maximum likelihood estimation finds θ by maximizing the likelihood $L(y|\theta, x)$ of observing y given x and θ and distributional assumptions about ε . A central assumption in this regular maximum likelihood estimation of the regression model is that x and ε are distributed independently, such that the joint distribution $f(x, \varepsilon)$ equals the product of the individual distributions $f(x)$ and $f(\varepsilon)$.

It is this latter assumption that the GC approach relaxes to obtain a modified maximum likelihood estimation, which in turn can easily be transformed into a modified OLS estimation. More specifically, the GC approach to addressing endogeneity starts with the observation that if the *joint* distribution $f(x, \varepsilon)$ of the regressor x and the error term ε were known, then it would be possible to obtain consistent estimates of the model parameters by maximizing the likelihood function derived from this joint distribution. Park and Gupta (2012) proposed using copulas to infer this joint distribution. Specifically, they made use of Sklar's (1959) theorem, which states that there exists a copula function, such that the joint cumulative distribution function (CDF) of x and ε can be derived from the marginal CDFs of x and ε . Furthermore, once the joint CDF of x and ε is known, calculating the joint probability distribution function $f(x, \varepsilon)$ is straightforward. Park and Gupta (2012) assumed the marginal distribution of ε to be normally distributed and derived the CDF of x based on observed realizations of x . They proposed using the Gaussian copula to link the two marginal distributions and thus derive the joint distribution $f(x, \varepsilon)$. Although this joint distribution appears complex at first, Park and Gupta (2012) showed that the Gaussian copula allows the maximization of the likelihood function to be rewritten in terms of an extended regression equation, which provides a simpler way to estimate the model:

$$y = \alpha + \beta x + \beta^c xc + \varepsilon^*, \quad (2)$$

with β^c another parameter to be estimated, ε^* normally distributed white noise and the additional regressor defined as $xc = \Phi^{-1}(F_x(x))$, where Φ^{-1} denotes the inverse of the standard normal CDF and F_x the CDF of x . Since F_x is unknown, the observed empirical CDF of x ,

\hat{F}_x , is used instead, and the extended regression equation becomes:

$$y = \alpha + \beta x + \beta^c \hat{x}\hat{c} + \varepsilon^*, \quad (3)$$

with $\hat{x}\hat{c} = \Phi^{-1}(\hat{F}_x(x))$. Least square regression then provides estimates of the parameters. Similar to the control function approach (Heckman, 1978), the GC approach thus simply adds an additional regressor $\hat{x}\hat{c}$ to the model. The parameter β^c is proportional to the correlation between x and ε , which has led previous research to interpret a significant β^c as indicative of endogeneity (e.g., Becerra & Markarian, 2021). Bootstrapping is required to obtain standard errors as the proposed method is a sequential approach that first estimates the density of the endogenous regressor and then augments the linear regression (Park & Gupta, 2012).

Review of GC Approach: Applications and Limitations

The introduction of the GC approach has made a strong impact on the marketing literature. Leading journals in marketing frequently publish papers that use the GC approach (e.g., Datta, Ailawadi, & van Heerde, 2017; Guitart & Stremersch, 2021). Similarly, the GC approach has seen a significant uptake in management research during the past years (e.g., Becerra & Markarian, 2021; Bhattacharya et al., 2020; Haschka & Herwartz, 2020; Reck et al., 2021). However, as with any statistical approach, its performance hinges on whether the underlying identifying assumptions are met; thus, a careful discussion of these assumptions is warranted.

When introducing the GC approach, Park and Gupta (2012) discussed its suitability in a range of situations and provided various simulations to test its robustness. For example, they showed how the GC approach can be used to deal with cases of multiple endogenous regressors and both discrete and continuous endogenous regressors. Park and Gupta (2012) also compared the GC approach against the Latent Instrumental Variable approach and found that it provides unbiased estimates even when misspecified, whereas the Latent Instrumental Variable estimator proved biased if its assumptions were not met. Finally, they tested the approach's performance when the distributional assumptions regarding the error term and the regressor are violated and concluded that the method is rather robust to these violations. While these robustness tests establish some confidence in the method, they do not necessarily match the peculiarities encountered in management research. Additionally, Becker et al. (2021) recently questioned the setup of these simulations and showed that results changed when accounting for an intercept in the estimation equation. The following discussion of the assumptions of the GC approach aims to shed more light on their importance for the performance of the method.

The first assumption of the GC approach is more implicit, as it is not immediately obvious in the derivation of the method. It states that the endogenous regressor needs to have sufficient support for the approach to be identified (Park & Gupta, 2012). As such, this assumption excludes situations where the endogenous regressor is binary or with few discrete outcomes, though Park and Gupta (2012) showed that their proposed method already is identified and provides unbiased estimates if the regressor has five discrete outcomes. This assumption makes the GC approach unsuitable in the case of binary regressors and distinctly less

useful in the case of other discrete regressors with limited support. As such, the GC approach should not be applied when dealing with endogeneity that is due to sample selection or self-selection into a limited number of outcomes (e.g., choice of alliance strategy). In the remainder of the paper, we focus solely on non-sample-selection-based endogeneity, and we use the term “endogeneity” to denote only endogeneity that is not induced by sample selection or self-selection.

The second assumption requires the exogenous regressor to be sufficiently nonnormally distributed. If the distribution of the endogenous regressor approaches the normal distribution, the correlation between x and \hat{x} increases, which in turn causes a multicollinearity problem and, in its extreme form, insufficient identification of parameters. If the regressor is not sufficiently nonnormally distributed, standard errors of the parameter estimates will be very large, leading researchers to fail to detect significant relationships. This nonidentification is demonstrated in Park and Gupta’s simulations in their online appendix, where—using t-distributions with $df=10$ and $df=30$, respectively, for the endogenous regressor—they found that “estimated values are in some instances far from the true values and consequently the SE’s of model parameters increase, resulting in false conclusion of insignificance” (Park & Gupta, 2012: online appendix V).

In general, Park and Gupta (2012) argued that, even in the case of normally distributed endogenous regressors, parameter estimates continue to be unbiased and are thus only a problem of efficiency, which can be mitigated by increasing the sample size. The importance of sample size is also highlighted in subsequent GC research (Becker et al., 2021; Falkenström, Park, & McIntosh, 2021), but many researchers may find this suggestion less helpful as the sample size is a frequent concern for applied research. Most management papers using the GC approach instead discuss the shape of the endogenous regressor (Becerra & Markarian, 2021; Bhattacharya et al., 2020). We also found one study that considered the GC approach but, based on the violation of this assumption, ended up not using it (Bitrián, Buil, & Catalán, 2021).

Given the wide variation of constructs and variables examined in management and business research, it is difficult to make a generalizing statement regarding whether this assumption is usually met for regressors in those fields of research. For example, the examination of firm performance (FP) variables by Certo, Raney, and Albader (2020) highlighted the nonnormality of many performance variables in strategic management. Similarly, variables like alliance activity, R&D spending, patent quality, and firm size have been shown to be nonnormal (e.g., Almeida, Hohberger, & Parada, 2011; Segarra & Teruel, 2012). On the other hand, there exist several examples of variables that are close to being normally distributed, including stock market returns measured as cumulative abnormal returns within event study frameworks (Sorescu, Warren, & Ertekin, 2017) and psychological variables like the intelligence quotient (Mendelson, 2000). As such, care should be taken when evaluating the appropriateness of the GC. Past applications of the GC approach frequently used the Shapiro-Wilk test (Shapiro & Wilk, 1965) to test the distribution of the endogenous regressor (Papies et al., 2016) and to determine whether the regressor is too close to normal to provide identification. However, the Shapiro-Wilk test is highly sensitive to large sample sizes and may reject the null hypotheses of normal distribution even if the departures from normality are trivial (Cortina & Dunlap, 1997). In fact, Becker et al. (2021) cautioned that normality tests are not suitable

to directly decide whether a distribution is sufficiently nonnormal for the GC approach to be identified.¹ We contribute to this discussion by comparing the GC's performance to its alternatives—namely the IV approach, which may similarly fail to detect significant relationships (Semadeni et al., 2014), and the HM approach (Ebbes et al., 2009), which requires asymmetric regressors for identification.

The third assumption of the GC approach requires the error term to be normally distributed. The difficulty of this assumption is that it cannot be tested directly. Although distributional tests or graphical inspections of the OLS residuals are usually used to get a sense of the error distribution, we note that under misspecification of the regression model, these residuals are not useful as they are then a function of the misspecification (McKean, Sheather, & Hettmansperger, 1993). Prior exploration of the error structure using uniform distribution of the error term (Park & Gupta, 2012) and nonnormal symmetric error distributions (Becker et al., 2021) provides mixed results on the robustness of the GC approach to this assumption. We advance this discussion by also analyzing the GC's performance in the context of skewed error terms as well as error terms that exhibit both skewness and positive kurtosis. Though researchers often argue that the errors, as the sum of many different unobserved factors, have an approximate normal distribution due to the central limit theorem, Wooldridge (2012) cautioned against this argument and stated that normality of the errors is an empirical matter. For example, many measures in strategic management exhibit nonnormal distributions with a high skew or a high kurtosis (Certo et al., 2020), which—depending on the model specification—may lead to nonnormal error terms. So far, only a small minority of management studies applying the GC approach have discussed the assumption regarding the normality of the error (e.g., Becerra & Markarian, 2021). Becerra and Markarian (2021) even transformed the variables of their model to obtain normally distributed residual, yet—as discussed above—normally distributed residuals in the case of model misspecification do not necessarily reflect normally distributed error terms.

The fourth and final assumption of the GC approach states that the dependency between the distribution of the endogenous regressor and the error term can be described by a Gaussian copula. Park and Gupta (2012) explored the method's robustness to violations of this assumption by showing that the method still performs well when the dependency structure is described by other copulas such as the Frank copula or the Plackett distribution, but these results were challenged by Becker et al. (2021), who found that the robustness of the GC approach in these cases was due to the estimation setup used by Park and Gupta (2012). Our simulations using a more agnostic data generation process also suggest that the violations of the GC dependency structure may have detrimental effects. Research in management using the GC approach has so far largely ignored this assumption. An exception to this is Bhattacharya et al. (2020), who mentioned that the GC approach is robust to violations of it.

Although tests exist to determine whether a particular copula describes the dependency structure between two variables (e.g., Huang & Prokhorov, 2014), these cannot be applied as the error term is not observable and—as outlined above—the residuals of the OLS are a function of the model misspecification. Thus, despite the importance of this assumption for the performance of the GC approach, researchers lack guidance on whether this assumption

may hold in their specific situation. We contribute to the field by discussing the link between the data-generating process that causes endogeneity in linear regression—namely reverse causality, measurement error, or omitted variables—and the dependency structure between error term and regressor.

It is worth noting that the different assumptions are interconnected. The requirement for the endogenous regressor to be sufficiently nonnormal while at the same time assuming that the error is normally distributed might pose a problem in empirical applications: Each of the three endogeneity sources leads to x and ϵ sharing a common random term, which in turn leads to them being more similar. For example, omitted variables with high skewness or high kurtosis that are highly correlated with x can lead to highly nonnormal error terms. We will return to this point in our discussion about the dependency structure between error terms and regressor.

Comparison to Other Instrument-Free Approaches

The GC approach is a relatively new addition to the family of instrument-free approaches for dealing with endogeneity. Instrument-free approaches aim to identify the model parameters through statistical assumptions about the error term, endogenous regressor, or functional form. Examples of instrument-free methods include the HM approach (Ebbes et al., 2009; Lewbel, 1997), the Identification through Heteroscedasticity estimator (Lewbel, 2012; Rigobon, 2003), and the Latent Instrumental Variables method (Ebbes et al., 2005; Sonnier, McAlister, & Rutz, 2011), to name a few. All of these instrument-free methods have in common that they rely heavily on distributional assumptions of the underlying variables (Lewbel, 2019). Instrument-free methods can in general be classified according to their estimation procedure as moment-based approaches or likelihood-based approaches (see Table 1).²

Moment-based approaches are generally more useful in large datasets because moment estimators can be erratic in small samples (Lewbel, 1997). These methods tend to be more susceptible to outliers but less sensitive to violations of the functional form. In contrast, likelihood-based approaches are less susceptible to outliers, but their performance depends on how well the data match the distributional assumptions of the method. An advantage of likelihood-based approaches over moment-based approaches is that they can be more easily adapted to more general situations than regular linear regression analysis.

Among the moment-based estimators, the HM approach was initially introduced by Lewbel (1997) to deal with endogeneity due to measurement error. This method creates artificial instruments using the available data and makes no demands on the distribution of the errors.³ Ebbes et al. (2009), however, showed that if symmetry of the error and nonsymmetry of the endogenous regressor is given, a variant of the HM approach can also be used to deal with endogeneity arising from omitted variable bias. Another moment-based estimator is the Identification through Heteroscedasticity estimator by Rigobon (2003), which requires heteroscedasticity of errors across different observable clusters of observations. Lewbel (2012) then showed how heteroscedasticity in general, combined with the existence of one truly exogenous regressor, can identify the parameter associated with the endogenous regressor.

Table 1
Comparison of Selected Instrument-Free Methods

IV free methods: Identification of model parameters in the presence of an endogenous regressor through statistical assumptions and functional form	
Estimation approach	<p>Moment based Researcher assumes functional form that relates parameter of interest to population moments of the dependent and independent variables. Sample moments are set equal to population moments to obtain estimate of parameter (Wooldridge, 2012: 768).</p> <p>Likelihood based Researcher makes assumptions about distribution of model variables and derives associated likelihood function. By maximizing this likelihood function the researcher obtains an estimate of the parameter of interest. (Wooldridge, 2012: 769).</p>
Method	<p>Higher Moments The researcher constructs instruments by exploiting higher order moments of the data (Ebbes et al., 2009; Erickson & Whited, 2002; Lewbel, 1997).</p> <p>Identification through heteroscedasticity The researcher constructs instruments by making assumptions about heteroscedasticity of errors (Lewbel, 2012; Rigobon, 2003).</p> <p>Latent Instrumental Variables The researcher assumes it is possible to linearly decompose the endogenous explanatory variable into categorical latent instruments and an iid error term (Ebbes et al., 2005; Sommer et al., 2011).</p> <p>Gaussian Copula The researcher assumes that the joint distribution of the endogenous regressor and the error term can be described by a Gaussian copula (Park & Gupta, 2012).</p>
Sources of Endogeneity & Main Assumptions	<p>Measurement error (Lewbel, 1997)</p> <ul style="list-style-type: none"> • Measurement and equation errors are independent and have moments of every order • Mismeasured regressor is asymmetric <p>Omitted variable (Ebbes et al., 2009)</p> <ul style="list-style-type: none"> • Error distribution is symmetrical • Endogenous regressor is asymmetric <p>Simultaneity, measurement error, or omitted variable (Lewbel, 2012)</p> <ul style="list-style-type: none"> • Existence of one exogenous regressor that is uncorrelated with errors • Errors heteroscedastic <p>Simultaneity (Rigobon, 2003)</p> <ul style="list-style-type: none"> • Heteroscedasticity of error across different observation groups • Needs information about group indicator <p>Simultaneity, measurement error, or omitted variable (Sommer et al., 2011)</p> <ul style="list-style-type: none"> • Error term is normally distributed • Endogenous regressor nonnormal • Existence of discrete latent instruments that describe regressor <p>Simultaneity, measurement error, or omitted variable (Rutz & Watson, 2019)</p> <ul style="list-style-type: none"> • Endogenous regressor has sufficient support • Gaussian Copula describes joint distribution of regressor and error term • Error term normal • Endogenous regressor nonnormal
Key Advantage	Can be adapted to more general situations than regular linear regression
Key Disadvantage	More sensitive to violations of distributional assumptions

The Latent Instrumental Variables method was proposed by Ebbes et al. (2005) as a likelihood-based alternative to achieve identification in the presence of endogeneity. The underlying assumption is that the endogenous regressor can be split into an exogenous and an endogenous part, with the former being described by at least two latent discrete instruments with different means and the latter being normally distributed. The challenges associated with justifying the assumptions of these different approaches, as well as the difficulties implementing them (Park & Gupta, 2012), are probably key reasons for the very limited use of Latent Instrumental Variables in business research (Hill et al., 2021).

Overall, instrument-free methods can be useful tools if their assumptions are met or if violations of their assumptions only cause minor deterioration in performance. In what follows we will compare the GC approach with the Ebbes et al. (2009) version of the HM approach, as the latter has similar underlying assumptions and is thus most likely to be used as an alternative to the GC approach if no suitable instruments are available.

Simulations

We conducted a series of Monte Carlo simulations to better understand the performance of the GC approach relative to OLS and IV regressions with instruments of different strengths and endogeneity as well as the HM approach developed for omitted variable bias and discussed in Ebbes et al. (2009). We used four sets of simulations to generate regressors and error terms that enable us to test the GC approach when (a) all assumptions of the GC approach are met, (b) the distribution of the regressor gets closer to normal, (c) the error term is not normally distributed, and (d) the dependency structure between the error and the endogenous regressor is described by a nonparametric data-generating mechanism. For each simulation set, we examined the impact of different levels of endogeneity and focused on effect sizes in the range that scholarship employs in simulations to ensure realistic effects (Busenbark et al., 2022; Certo et al., 2016; Certo et al., 2020; Semadeni et al., 2014). All simulations were based on 500 runs with $N = 500$ observations each and were performed in R version 4.02.

To compare the suitability of different estimation methods, researchers often use the concepts of bias, precision, and power (Semadeni et al., 2014). Bias occurs when the expected value of the estimator and the true value of the parameter diverge. In other words, an unbiased estimation method applied to several datasets generated by the same underlying true parameter will lead to parameter estimates with a mean equaling the true underlying parameter, whereas a biased estimation method will not. We use two indicators to measure bias: first, the median estimate of β (*beta*), where values of the median β that exceed (are less than) their true values suggest positive (negative) bias, and second, the t-bias (*t-bias*) measure, which measures the absolute deviation of the mean of the sampling distribution from the true parameter value expressed in terms of number of standard errors of the parameter estimate (see also Park & Gupta, 2012). The t-bias reflects the Type 1 error of the models (i.e., the incorrect rejection of a true null hypothesis). If the t-bias is smaller than 2 (corresponding to the critical t-statistic for a test at 95% confidence), the confidence interval of the estimate will include the true value.

Precision refers to the spread of the estimator if applied multiple times. For two unbiased estimates, the one with higher precision (measured by a lower standard error) should be

preferred, as this estimator will more likely reflect the true underlying parameters. We account for precision via the median standard error (*SE*) for the 500 simulation runs.

Power is a method's ability to reject a null hypothesis of no effect when it is false. Based on the discussion in Park and Gupta (2012) regarding the possible inability of the GC approach to detect significant relationships when the regressor is not sufficiently nonnormally distributed, we give particular emphasis to this issue. Thus, we measure the percentage of significant estimates (*% sign beta*), which reports the percentage of estimations that result in statistically significant coefficients. This measure illustrates each method's susceptibility to Type 2 errors (i.e., to incorrectly retain a false null hypothesis). We also report the percentage of times the correction term in the GC approach is significant at $p = 0.1$ (*% sign beta^c*) to test whether its significance is reflective of endogeneity in the regressor. Some studies have used this approach to argue that endogeneity is present (e.g., Datta et al., 2017), while others have raised concerns about the power of this estimate (e.g., Becker et al., 2021). Thus, a careful investigation of the associated Type 1 and Type 2 error of this estimate is warranted to explore the suitability of this test to confirm or dismiss the existence of endogeneity.

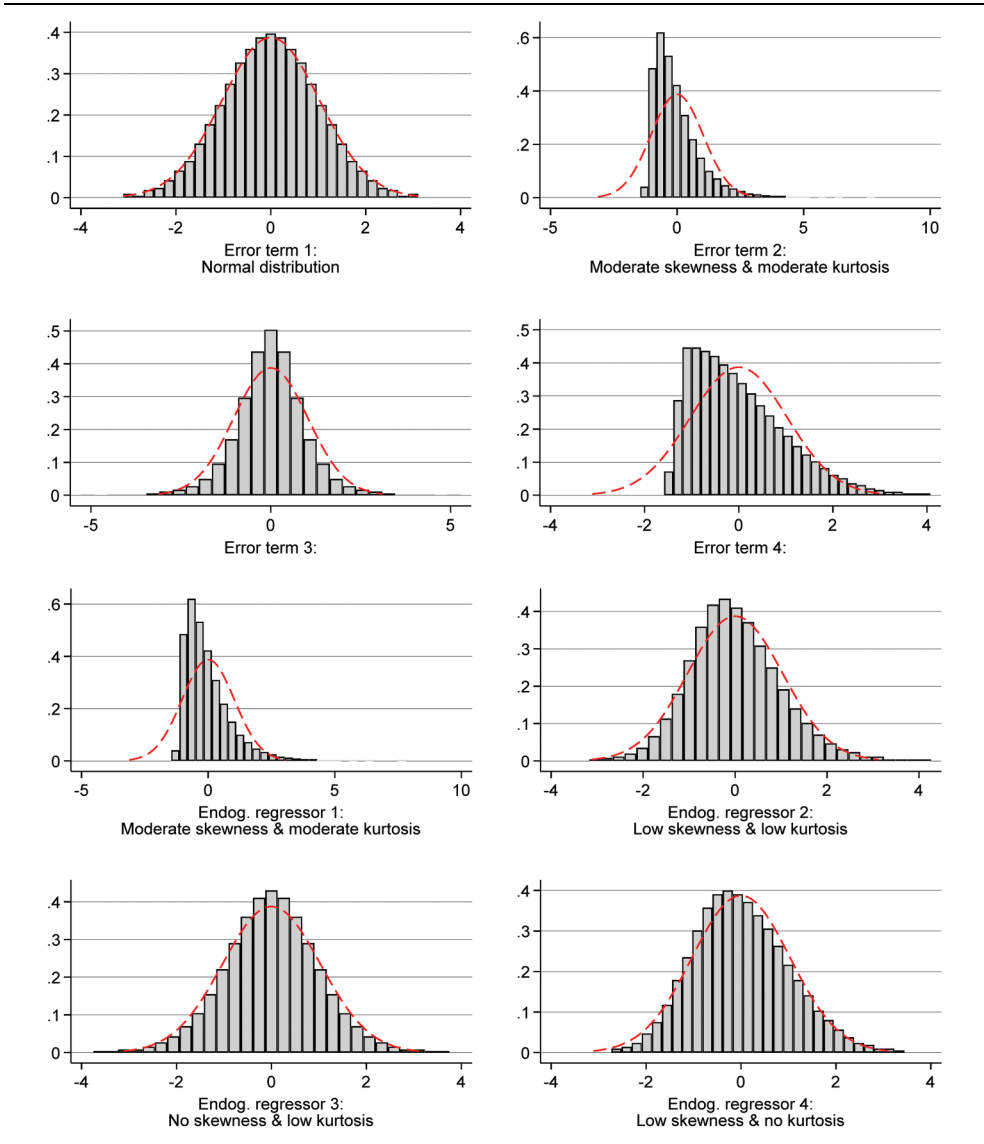
In all of the simulations, we generate a dependent variable y via $y = \beta x + \varepsilon$. In Simulation 1 to Simulation 3, we define $x = f_x(x^*)$ and $\varepsilon = g_\varepsilon(\varepsilon^*)$ with x^* and ε^* drawn from the bivariate normal distribution with a prespecified correlation, which ensures that x^* and ε^* follow a Gaussian copula. In contrast, the monotonic functions f_x and g_ε transform these draws so that the resulting regressors x and error terms ε in the simulations have the desired distributions. Further, since copulas are invariant under monotonic transformations (Nelsen, 2007), they also ensure that the regressors x and error terms ε follow a Gaussian copula. In Simulation 4, we generate x and ε directly using nonparametric sampling. This type of sampling allows specifying desired distributions for x and ε as well as desired correlations without the need for any assumptions about the underlying copula structure.

Simulation 1: Baseline Simulation When All Model Assumptions Are Met

We begin by reviewing the performance of the GC approach when all of its model assumptions are met. We assume regressors with moderate skewness and moderate kurtosis as well as errors that are normally distributed. The different distributions used for the error terms and endogenous regressors (including skewness and kurtosis values) are shown in Figure 1.

The simulation setup of this case will be the blueprint for the following simulations, where specific assumptions are varied and tested; it is described in the following steps. First, we define the desired correlation ρ between regressor and error, $\rho \in \{0, 0.1, 0.3\}$, representing no, low, and medium endogeneity conditions, respectively. Then we define the desired parameter $\beta \in \{0, 0.1, 0.3\}$. Since we standardize our regressor x and the dependent variable y , these parameter values translate into no, small, and medium effects of the regressor on the dependent variable (Cohen, 1988). Next, we repeat the following steps until we have obtained 500 full simulation runs: (Step 1) We generate $N = 500$ bivariate random normal variables (x^*, ε^*) with the desired correlation. (Step 2) We apply the following monotonic transformations $f_x: \mathbb{R} \rightarrow \mathbb{R}$, $g_\varepsilon: \mathbb{R} \rightarrow \mathbb{R}$ to x^* and ε^* : $x = f_x(x^*) = f(\Phi(x^*))$, $\varepsilon = g_\varepsilon(\varepsilon^*) = g(\Phi(\varepsilon^*))$, where Φ is the cumulative density function of the standard normal distribution, f is the quantile function of the skewed generalized t distribution ($f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.8, p = 2, q = 2)$), and g is the quantile function of the standard

Figure 1
Distribution of the error terms and endogenous regressors



normal distribution (as such, $\varepsilon = \varepsilon^*$ in this set of simulations). We then standardize x and rescale ε so that the resulting variable y in the next step has a variance of 1. Since the correlation between two variables can be distorted by nonlinear transformations, we discard any simulation runs that do not yield the desired correlation.⁴ We also discard any simulation runs where the Shapiro-Wilk test did not reject the null hypothesis of normal distribution

Table 2
Baseline Simulation When All Model Assumptions Are Met

	No endog. $\text{cor}(x,e)=0$				Low endog. $\text{cor}(x,e)=0.1$				Medium endog. $\text{cor}(x,e)=0.3$			
	Median beta	Median S.E.	t-bias beta	% sign beta ^c	Median beta	Median SE	t-bias beta	% sign beta ^c	Median beta	Median S.E.	t-bias beta	% sign beta ^c
<i>Panel A: Effect size = 0.0</i>												
OLS	0.002	0.045	0.638	6.0%	0.100	0.045	2.240	100.0%	0.298	0.043	6.972	100.0%
IV (exog; weak)	0.008	0.448	0.606	0.2%	0.012	0.506	0.582	0.4%	0.051	0.479	0.596	2.6%
IV (exog; medium)	0.002	0.148	0.641	1.2%	0.003	0.153	0.623	3.4%	0.009	0.150	0.658	4.8%
IV (exog; strong)	0.001	0.089	0.586	1.2%	0.002	0.090	0.568	2.0%	0.005	0.090	0.590	3.0%
IV (endog; weak)	0.979	0.648	1.391	13.4%	1.074	0.640	1.440	18.2%	1.004	0.562	1.734	37.2%
IV (endog; medium)	0.340	0.159	2.154	60.0%	0.341	0.155	2.218	62.2%	0.340	0.145	2.336	66.8%
IV (endog; strong)	0.202	0.092	2.220	62.8%	0.205	0.090	2.277	64.0%	0.198	0.086	2.325	63.0%
GC	0.004	0.111	0.566	5.6%	0.021	0.110	0.668	2.2%	0.017	0.104	0.743	5.6%
HM	-0.001	0.059	0.768	10.0%	0.066	0.060	1.118	20.0%	0.179	0.059	3.073	83.6%
<i>Panel B: Effect size = 0.1</i>												
OLS	0.096	0.045	0.678	59.0%	0.199	0.044	2.238	100.0%	0.393	0.042	6.999	100.0%
IV (exog; weak)	0.072	0.504	0.618	1.6%	0.086	0.492	0.589	1.2%	0.081	0.492	0.605	2.8%
IV (exog; medium)	0.097	0.152	0.662	9.4%	0.095	0.151	0.639	8.2%	0.092	0.149	0.650	10.4%
IV (exog; strong)	0.098	0.090	0.603	17.8%	0.097	0.089	0.582	16.6%	0.096	0.088	0.583	20.6%
IV (endog; weak)	1.073	0.658	1.357	26.0%	1.099	0.623	1.432	30.4%	1.071	0.547	1.663	43.8%
	0.425	0.159	2.082	83.4%	0.424	0.154	2.151	83.2%	0.424	0.143	2.278	84.2%

(continued)

Table 2 (continued)

	No endog. cor(x,e) = 0				Low endog. cor(x,e) = 0.1				Medium endog. cor(x,e) = 0.3			
	Median beta	Median S.E.	t-bias	% sign beta^c	Median beta	Median SE	t-bias	% sign beta^c	Median beta	Median S.E.	t-bias	% sign beta^c
IV (endog.; medium)	0.296	0.091	2.164	95.6%	0.296	0.089	2.204	95.2%	0.299	0.084	2.381	96.0%
IV (endog.; strong)	0.093	0.111	0.572	13.0%	0.117	0.109	0.606	17.6%	0.114	0.103	0.595	19.8%
GC	0.093	0.061	0.688	32.6%	0.165	0.059	1.126	78.6%	0.278	0.058	3.161	97.2%
HM												
<i>Panel C: Effect size = 0.3</i>												
OLS	0.298	0.043	0.689	100.0%	0.394	0.042	2.247	100.0%	0.573	0.039	7.006	100.0%
IV (exog.; weak)	0.318	0.450	0.614	6.0%	0.289	0.470	0.547	7.6%	0.344	0.455	0.575	14.8%
IV (exog.; medium)	0.307	0.143	0.654	57.6%	0.296	0.143	0.572	54.2%	0.313	0.139	0.607	62.2%
IV (exog.; strong)	0.303	0.086	0.593	95.8%	0.298	0.085	0.523	97.0%	0.307	0.082	0.553	97.0%
IV (endog.; weak)	1.264	0.617	1.436	48.4%	1.230	0.594	1.473	51.2%	1.254	0.494	1.689	63.4%
IV (endog.; medium)	0.623	0.152	2.145	99.0%	0.612	0.146	2.152	99.6%	0.610	0.133	2.438	99.4%
IV (endog.; strong)	0.495	0.088	2.224	100.0%	0.485	0.085	2.210	100.0%	0.477	0.078	2.235	100.0%
GC	0.292	0.104	0.576	78.8%	0.315	0.102	0.619	83.6%	0.313	0.095	0.624	88.4%
HM	0.295	0.057	0.685	99.4%	0.360	0.056	1.038	1.000	0.464	0.052	3.211	100.0%

Note: Endog. Regressor 5 = moderate skewness and moderate kurtosis; error distribution 1 = normal distribution.

at $p = 0.1$ for the regressor x to mimic the prevalent screening-pattern researchers employ to rule out normality in the context of the GC approach (Papies et al., 2016). (Step 3) We generate $y = \beta x + e$. (Step 4) We generate instruments $iv = \mu + \gamma x + \delta e$, where μ are random draws that have the same distribution as x . Similar to Semadeni et al. (2014) and Larcker and Rusticus (2010), we set γ and δ in such a way that the instruments have the desired strength ($cor(x, iv) \in \{0.1, 0.3, 0.5\}$) and endogeneity ($cor(e, iv) \in \{0, 0.1\}$) at the population level (see Online Appendix A for calculation of γ and δ). (Step 5) We use OLS, IV regressions, GC and HM approach to obtain parameter estimates for the regression of y on x .⁵

Table 2 shows the results of this initial simulation. The medians of the estimates obtained via the GC approach are very close to the true parameter value. The difference between the true parameter and the mean of the estimated parameters is usually smaller for the GC than the HM approach when endogeneity is present. The GC approach thereby provides estimates that are in most cases slightly closer to the true value than those obtained from regressions with exogenous, weak instruments. The median standard errors of the GC estimates are in all cases smaller than the median standard errors of exogenous instruments of medium strength yet larger than the median standard errors of the HM approach. While smaller standard errors are generally desirable, the t-bias statistic shows that they create difficulties in the HM approach. For the medium endogeneity scenarios, the HM estimates are on average more than 2 SEs away from the true parameters, implying that the HM estimates are biased. In contrast, the GC approach produces estimates with t-bias values smaller than 0.75 across all endogeneity conditions. This t-bias is larger than the one reported in Park and Gupta (2012). Yet in line with the bias reported in Becker et al. (2021, Simulation Study 1), who show that the inclusion of an intercept in the estimation equation leads to a deterioration of the GC's performance. We note that in the case of no effect ($\beta = 0$), the GC approach only detects significant relationships in less than 6% of cases. In the conditions of small and medium effect sizes, the GC approach's ability to detect significant relationships surpasses that of IV regressions with medium strength, exogenous instruments. In the no-endogeneity condition with nonzero effect sizes, IV regressions with exogenous instruments, the GC approach, and the HM approach all detect far less significant relationships than the correctly specified OLS. This finding is in line with previous observations regarding the general inefficiency of the IV approach (Murray, 2006) and instrument-free approaches (Lewbel, 2012).

We conclude by observing that the share of simulations in which the GC correction parameter is significant is below 10% in the case of no endogeneity, hovers around 23% in the case of low endogeneity, and increases to 90% in the case of medium endogeneity. This implies that the significance of this parameter is indicative of endogeneity, but it should not be used to argue for endogeneity nor to claim that endogeneity is not sufficient to bias OLS estimates.

Simulation 2: Simulation Results for Different Distributions of the Endogenous Regressor

Next, we focus on the performance of the GC approach in the context of regressors that approach the normal distribution. We continue to generate the errors as normally distributed and maintain the GC assumption to generate the joint distribution between regressor and error. To generate the data, we use the same setup as outlined before but choose different distributions for the regressor x . More specifically, in Step 2 we use for f the quantile functions of

Table 3
Simulation Results for Different Distribution of the Endogenous Regressor

		No endog. $\text{cor}(x,e) = 0$				Low endog. $\text{cor}(x,e) = 0.1$				Medium endog. $\text{cor}(x,e) = 0.3$					
		Median beta	SE	t-bias	% sign beta	% sign β^2	Median beta	SE	t-bias	% sign beta	% sign β^2	Median S.E.	t-bias	% sign beta	% sign β^2
<i>Panel A: Endog. regressor 2, low skewness and low kurtosis</i>															
OLS		0.097	0.045	0.715	59.2%		0.199	0.044	2.243	100.0%		0.393	0.042	6.995	100.0%
IV (exog; weak)		0.090	0.462	0.589	1.0%		0.109	0.486	0.561	1.8%		0.088	0.462	0.586	4.6%
IV (exog; medium)		0.098	0.149	0.634	8.6%		0.103	0.150	0.618	8.8%		0.094	0.147	0.642	12.0%
IV (exog; strong)		0.099	0.089	0.571	14.8%		0.102	0.089	0.558	17.2%		0.097	0.088	0.577	18.8%
IV (endog; weak)		1.023	0.609	1.402	26.2%		1.143	0.598	1.496	32.4%		1.105	0.533	1.695	46.2%
IV (endog; medium)		0.424	0.155	2.107	84.0%		0.435	0.151	2.203	84.4%		0.429	0.140	2.290	86.4%
IV (endog; strong)		0.296	0.091	2.157	94.8%		0.300	0.089	2.248	96.2%		0.295	0.085	2.356	94.6%
GC		0.086	0.385	0.646	5.4%	8.0%	0.125	0.375	0.573	3.8%	7.6%	0.175	0.365	0.636	8.0%
HM		0.093	0.197	0.521	8.4%		0.167	0.187	0.554	13.0%		0.295	0.187	1.088	37.8%
<i>Panel B: Endog. regressor 3, no skewness and low kurtosis</i>															
OLS		0.104	0.045	0.714	63.8%		0.199	0.044	2.242	100.0%		0.393	0.042	7.000	100.0%
IV (exog; weak)		0.113	0.476	0.590	1.0%		0.128	0.473	0.614	1.8%		0.075	0.477	0.605	3.6%
IV (exog; medium)		0.104	0.150	0.633	7.8%		0.106	0.149	0.658	8.4%		0.088	0.148	0.658	11.4%
IV (exog; strong)		0.102	0.089	0.574	16.2%		0.104	0.089	0.594	17.4%		0.093	0.088	0.594	16.8%
		1.081	0.631	1.421	25.8%		1.039	0.585	1.514	34.2%		1.018	0.515	1.655	44.8%

(continued)

Table 3 (continued)

	No endog. cor(x,e) = 0				Low endog. cor(x,e) = 0.1				Medium endog. cor(x,e) = 0.3						
	Median beta	SE	t-bias	% sign beta	% sign beta^c	Median beta	SE	t-bias	% sign beta	% sign beta^c	Median beta	SE	t-bias	% sign beta	% sign beta^c
	IV (endog.; weak)	0.435	0.157	2.128	82.2%		0.430	0.152	2.220	83.0%		0.413	0.141	2.248	82.8%
IV (endog.; medium)	0.301	0.091	2.196	94.8%		0.299	0.089	2.271	95.0%		0.304	0.085	2.419	94.4%	
IV (endog.; strong)					7.4%					7.0%					9.0%
GC	0.111	0.459	0.614	4.4%		0.162	0.455	0.588	6.6%		0.211	0.426	0.652	4.8%	
HM	0.145	0.510	0.481	2.6%		0.178	0.500	0.540	5.4%		0.360	0.452	0.727	11.0%	
<i>Panel C: Endog. regressor 4, low skewness and no kurtosis</i>															
OLS	0.102	0.045	0.648	63.6%		0.199	0.044	2.240	100.0%		0.393	0.042	7.001	100.0%	
IV (exog.; weak)	0.113	0.497	0.622	1.0%		0.060	0.477	0.590	2.4%		0.093	0.475	0.605	3.4%	
IV (exog.; medium)	0.104	0.152	0.688	10.0%		0.088	0.149	0.653	8.0%		0.090	0.148	0.704	11.4%	
IV (exog.; strong)	0.102	0.090	0.629	20.4%		0.094	0.089	0.596	16.2%		0.094	0.088	0.639	19.2%	
IV (endog.; weak)	1.143	0.667	1.399	25.2%		1.062	0.581	1.453	32.0%		1.017	0.532	1.626	43.2%	
IV (endog.; medium)	0.444	0.160	2.154	82.8%		0.420	0.151	2.109	82.4%		0.415	0.141	2.258	81.2%	
IV (endog.; strong)	0.306	0.092	2.220	92.4%		0.291	0.089	2.160	94.2%		0.302	0.084	2.411	95.8%	
GC	0.127	0.523	0.609	4.4%	7.8%	0.135	0.529	0.665	6.0%	8.2%	0.226	0.505	0.641	4.4%	8.4%
HM	0.099	0.207	0.529	6.4%		0.171	0.212	0.560	13.4%		0.310	0.204	1.034	36.0%	

Note: Error distribution 1 = normally distributed, effect size = 0.1.

the skewed generalized t distributions $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.2, p = 2, q = 5)$, $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0, p = 2, q = 5)$, and $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.2, p = 2, q = INF)$, which results in regressors of low skew and low kurtosis, no skew and low kurtosis, and low skew and no kurtosis, respectively. We present the results obtained for a small effect size ($\beta = 0.1$) and note that results remain fundamentally unchanged if other effect sizes are employed.

The results in Table 3 reveal that despite all of the regressor distributions being very close to the normal distribution, the median of the t -bias statistic of the GC approach is smaller than 0.67, indicating that the GC approach yields parameters that are less than 1 SE away from the true parameter. Nonetheless, the medians of the parameter estimates are farther away from the true parameter the higher the endogeneity of the regressor, with median estimates that go up to more than double the size of the original estimate. In the presence of endogeneity, the estimated parameters are closer to the true values than those obtained with OLS, IV regressions with strong but endogenous instruments, or the HM approach. Yet they are mostly farther away from the true values than those obtained with weak but exogenous instruments.⁶ Given that all our simulations are based on regressors for which the Shapiro-Wilk test did reject the assumption of normality at $p < 0.1$, this statistic does not seem a strong enough test to exclude the possibility that the regressor deviates enough for the GC approach to be identified.⁷

We find that the size of the standard errors of the GC approach usually lies between weak and medium exogenous instruments and so does its ability to detect significant relationships. Across all conditions, it only finds significant effects in less than 10% of simulations. It thereby performs consistently worse than the HM approach, which—also because of its upwardly biased estimates—detects significant relationships more frequently. The significance of the GC correction parameter seems rather uninformative in these simulations: The parameter is only significant in a maximum of 12% of cases across the different conditions.

Simulation 3: Simulation Results for Different Error Distributions

In this set of simulations, we investigate the performance of the GC approach when the error terms are not normally distributed. More specifically, in Step 2 we change g and use the quantile functions of the skewed generalized t distributions $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.8, p = 2, q = 2)$, $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0, p = 2, q = 2)$, and $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.8, p = 2, q = INF)$, which results in error terms of moderate skew and moderate kurtosis, no skew and moderate kurtosis, and moderate skew and no kurtosis, respectively. We generate the regressor with moderate skew and moderate kurtosis [so f is the quantile function of the skewed generalized t distribution $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.8, p = 2, q = 2)$] and maintain the GC assumption to generate the joint distribution between regressor and error. Again, we focus on the case of a small effect size ($\beta = 0.1$).⁸

Table 4 shows that the median of the parameter estimates of the GC approach is close to the true parameters in no- and low-endogeneity conditions or when the error terms exhibit no skew, yet the estimated values reach up to double the size of the true value when the error term exhibits skew and the correlation between regressor and error term is of medium strength. While the GC approach performs worse than IV regression with weak instruments in these

Table 4
Simulations With Varying Distributions of the Error Term

	No endog. $\text{cor}(x,e)=0$				Low endog. $\text{cor}(x,e)=0.1$				Medium endog. $\text{cor}(x,e)=0.3$			
	Median	Median	% sign	% sign	Median	Median	% sign	% sign	Median	Median	% sign	% sign
	beta	S.E.	t-bias	beta ² c	beta	SE	beta	beta ² c	beta	SE	beta	beta ² c
<i>Panel A: Error term 2, Moderate skewness and moderate kurtosis</i>												
OLS	0.096	0.045	0.707	58.6%	0.199	0.044	2.245	100.0%	0.394	0.042	7.009	100.0%
IV (exog.; weak)	0.045	0.479	0.598	1.0%	0.068	0.497	0.578	1.2%	0.080	0.475	0.544	5.2%
IV (exog.; medium)	0.085	0.150	0.630	6.6%	0.087	0.151	0.607	7.8%	0.087	0.146	0.627	11.8%
IV (exog.; strong)	0.092	0.089	0.573	14.0%	0.093	0.090	0.553	14.6%	0.093	0.088	0.564	17.4%
IV (endog.; weak)	1.016	0.617	1.338	22.4%	1.079	0.621	1.441	29.6%	1.091	0.533	1.701	46.0%
IV (endog.; medium)	0.416	0.156	2.026	81.8%	0.420	0.154	2.104	83.2%	0.423	0.142	2.246	85.2%
IV (endog.; strong)	0.291	0.091	2.112	95.0%	0.292	0.089	2.151	95.6%	0.299	0.084	2.397	93.4%
GC	0.088	0.100	0.654	11.4%	0.114	0.111	0.556	10.0%	0.218	0.128	0.900	32.2%
HM	0.093	0.070	0.712	22.6%	0.203	0.083	1.217	0.786	0.410	0.098	3.173	98.8%
<i>Panel B: Error term 2, No skewness and moderate kurtosis</i>												
OLS	0.100	0.045	0.661	61.4%	0.199	0.044	2.237	100.0%	0.393	0.042	6.984	100.0%
IV (exog.; weak)	0.098	0.497	0.575	0.6%	0.099	0.458	0.645	1.4%	0.133	0.484	0.620	4.2%
IV (exog.; medium)	0.099	0.152	0.599	8.0%	0.097	0.148	0.705	9.0%	0.109	0.148	0.676	11.4%
IV (exog.; strong)	0.100	0.090	0.544	14.6%	0.098	0.089	0.637	18.0%	0.105	0.088	0.612	20.6%
	1.102	0.658	1.384	25.2%	1.012	0.579	1.460	32.6%	1.126	0.541	1.676	43.8%

(continued)

Table 4 (continued)

	No endog. cor(x,e)=0				Low endog. cor(x,e)=0.1				Medium endog. cor(x,e)=0.3			
	Median beta	Median S.E.	t-bias	% sign beta [^] c	Median beta	Median SE	t-bias	% sign beta [^] c	Median beta	Median SE	t-bias	% sign beta [^] c
IV (endog.; weak)	0.427	0.159	2.096	83.2%	0.422	0.151	2.155	82.8%	0.434	0.142	2.401	85.0%
IV (endog.; medium)	0.300	0.091	2.181	95.2%	0.295	0.089	2.218	93.8%	0.290	0.084	2.267	94.2%
IV (endog.; strong)	0.099	0.107	0.595	14.8%	0.128	0.105	0.689	22.0%	0.106	0.110	0.613	13.6%
GC	0.099	0.059	0.715	38.6%	0.166	0.059	1.094	78.4%	0.286	0.067	2.652	98.0%
HM												
<i>Panel C: Error term 2, Moderate skewness and no kurtosis</i>												
OLS	0.095	0.045	0.718	56.8%	0.199	0.044	2.240	100.0%	0.393	0.042	6.983	100.0%
IV (exog.; weak)	0.143	0.498	0.601	2.4%	0.110	0.488	0.518	1.8%	0.064	0.466	0.556	3.8%
IV (exog.; medium)	0.114	0.151	0.634	10.4%	0.103	0.150	0.565	8.6%	0.087	0.147	0.614	9.6%
IV (exog.; strong)	0.107	0.090	0.577	19.6%	0.102	0.089	0.514	17.8%	0.093	0.088	0.554	16.8%
IV (endog.; weak)	1.138	0.672	1.404	24.8%	1.101	0.624	1.509	35.4%	1.092	0.525	1.668	44.8%
IV (endog.; medium)	0.450	0.158	2.184	84.0%	0.426	0.153	2.197	84.4%	0.423	0.141	2.242	87.0%
IV (endog.; strong)	0.308	0.091	2.261	96.0%	0.297	0.089	2.258	96.0%	0.300	0.084	2.413	95.4%
GC	0.088	0.108	0.686	12.0%	0.127	0.115	0.636	16.6%	0.161	0.113	0.728	29.2%
HM	0.090	0.069	0.705	24.6%	0.198	0.077	1.278	81.4%	0.377	0.081	3.412	98.6%

Note: Endog. Regressor 1 = moderate skewness & moderate kurtosis, effect size = 0.1.

situations, it again performs better than OLS, IV regressions with endogenous but strong instruments, or the HM approach. And in contrast to these, the GC approach provides estimates in all conditions that are less than 1 SE away from the true values as indicated by the t-bias. Although this is due to its larger standard errors, these standard errors are reasonably tight compared to IV regressions. The median of the standard errors of the parameter estimates of the GC approach is for all conditions between those of medium and strong exogenous instruments, and—combined with the upward bias of the GC estimates—it allows the GC approach to detect significant relationships in more simulations than IV regressions with exogenous instruments of medium strength. Finally, in contrast to the prior simulations where the regressor was close to normal, the GC correction parameter is now again better able to pick up endogeneity. It detects endogeneity in up to 26% of simulations with low endogeneity and approximately 60% of simulations with medium endogeneity, with this number being higher when the error term is not skewed. Yet this parameter is also significant in up to 10% of cases when no endogeneity is present.

Simulation 4: Simulations to Test the Dependency Structure Between Error and Endogenous Regressor

Finally, we present the results when the dependency structure between endogenous regressor x and error term ϵ cannot be described by a Gaussian copula. We maintain the other assumptions of the GC approach by generating regressors with moderate skew and moderate kurtosis [so f is the quantile function of the skewed generalized t distribution $f_{SGT}(\mu = 0, \sigma = 1, \lambda = 0.8, p = 2, q = 2)$], normally distributed error terms, and a small effect size ($\beta = 0.1$).

There exists an infinite number of possible dependency structures that may generate the desired correlation between x and ϵ . Rather than arbitrarily picking specific ones, we generate data without assuming any underlying statistical model or copula that describes the dependency structure (Ruscio & Kaczetow, 2008). Therefore, we replace Steps 1 and 2 of the simulations with a nonparametric sampling of random variables of the desired distributions and the desired correlation using R's package SimJoint.

Table 5 details the results of this simulation. It reveals that the GC approach provides estimates that are far away from the true values in any endogeneity condition and that are significantly biased in the presence of endogeneity as indicated by a t-bias statistic larger than 2. In all cases, the GC's estimates are farther away from the true value than estimates obtained via OLS. While it shares this fate with the HM approach or endogenous instruments, its estimates are even farther away from the true values. We also note that the GC correction parameter is significant in approximately 10% of cases, and it does not distinguish between the different levels of endogeneity used in the simulation.

Discussion of Simulation Results

Our simulation studies provide new and relevant insights into the absolute and relative performance of the GC approach. We show that GC performs almost on par with IV regressions with exogenous instruments of medium strength when all its assumptions are met. It also provides parameter estimates that are closer to the true values than those obtained with OLS or

Table 5
Nonparametric Dependency Structure Between Error and Endogenous Regressor

	No endog. $\text{cor}(x,e) = 0$				Low endog. $\text{cor}(x,e) = 0.1$				Medium endog. $\text{cor}(x,e) = 0.3$				
	Median beta	Median SE	t-bias beta	% sign beta^c	Median beta	Median S.E.	t-bias beta	% sign beta^c	Median beta	Median SE	t-bias beta	% sign beta	% sign beta^c
OLS	0.100	0.045	0.000	100.0%	0.199	0.044	2.243	100.0%	0.394	0.042	7.018	100.0%	
IV (exog.; weak)	0.141	0.502	0.589	0.8%	0.096	0.469	0.594	1.8%	0.070	0.473	0.525	4.4%	
IV (exog.; medium)	0.113	0.152	0.637	8.6%	0.099	0.149	0.655	9.4%	0.087	0.148	0.586	9.6%	
IV (exog.; strong)	0.107	0.090	0.581	18.2%	0.099	0.089	0.596	18.2%	0.093	0.088	0.529	15.0%	
IV (endog.; weak)	1.113	0.657	1.401	20.8%	1.029	0.559	1.491	33.2%	1.085	0.528	1.597	43.2%	
IV (endog.; medium)	0.438	0.159	2.180	82.2%	0.429	0.151	2.176	85.6%	0.423	0.142	2.242	83.8%	
IV (endog.; strong)	0.303	0.091	2.267	95.0%	0.299	0.089	2.245	96.2%	0.300	0.084	2.413	94.0%	
GC	0.223	0.108	1.176	76.4%	0.325	0.106	2.188	97.4%	0.510	0.100	4.204	98.4%	7.8%
HM	0.110	0.041	0.243	98.2%	0.210	0.041	2.699	100.0%	0.404	0.039	7.773	99.2%	

the HM approach if the endogenous regressor is very close to normal or the error terms are not normally distributed, yet it performs worse than IV regressions with exogenous instruments.

We also show that the GC's ability to recover the true parameters is most strongly affected by the assumption that has received very little attention in past applications, namely, the assumption that the dependency structure between regressor and error term can be described by a GC. In our simulations, we used an agnostic, nonparametric approach to generating our simulated data that does not require assumptions about an underlying statistical model and that probably presents one of the more severe violations of the GC approach's assumption compared to evaluations in prior studies (Becker et al., 2021; Park & Gupta, 2012).

Even though the dependency structure between the regressor and the error term is unobservable, we suggest that considering the different sources of endogeneity and their effect on the relationship between regressor and error term may provide some guidance on whether the assumption regarding the Gaussian copula is met. More specifically, under certain assumptions, these endogeneity types can naturally lead to the GC dependency structure. First, for all sources of endogeneity, the endogenous regressor and the error term can be written as a linear combination of independent random variables (see Online Appendix B). Second, if we assume that these random variables are normally distributed, then we know that the endogenous regressor and the error term are also jointly normally distributed. Further, since the Gaussian copula produces the bivariate normal distribution and since Sklar's theorem tells us that this copula is unique (Durante, Fernandez-Sanchez, & Sempì, 2013), we know that the Gaussian copula describes the dependency structure between the endogenous regressor and the error term. Third, while these normality assumptions regarding the random variables result in a normally distributed endogenous regressor, it also holds that copulas are invariant under monotonic transformations (Nelsen, 2007)—that is, any strictly increasing transformation of the regressor will still be joined with the error term via a Gaussian copula.

As such, the GC dependency structure can be appropriate in a variety of endogeneity settings. Similarly, researchers can find further reassurance in the appropriateness of the GC assumption by referring to Danaher and Smith (2011), who reported that the Gaussian copula is a robust copula for most applications.

We conclude this section by noting that while our simulations demonstrate the performance of the different approaches in a wide range of settings, they are naturally limited in scope. For example, our choice of simulation setup might impact the comparison of the GC and HM approach, as we did not include different sample sizes. Further, although our results indicate that the GC approach performs well as long as the error term exhibits no skew, Becker et al. (2021) found that positive or negative kurtosis may also impact the GC's performance. We also did not generate endogeneity as resulting from omitted variable bias, which is the data-generation mechanism assumed in the HM version by Ebbes et al. (2009). As such, the performance of the GC approach relative to the HM approach in our simulations should only be used as an indication of its relative performance in general.

Empirical Examples

To demonstrate the application of the GC approach in a management context, we replicate two empirical studies in two distinct settings: (1) Deng, Kang, and Low (2013), with a focus on the relationship between CSR and FP, and (2) Autor, Dorn, Hanson, Pisano, and Shu

(2020), with a focus on the relationship between international trade and innovation. Both studies applied OLS and IV regression, thus providing us with a benchmark of comparison for the GC approach. An important distinction between the studies is the distribution of the endogenous regressor, which is highly nonnormal in Autor et al. (2020) but closer to normal in Deng et al. (2013). This difference allows us to examine the importance of the GC approach's only testable assumption of nonnormality of the regressor and its sensitivity to possible violations in an empirical context. We start with a "narrow" replication (Bettis, Helfat, & Shaver, 2016) using the data provided by the authors.⁹ We focus on the central argument or hypothesis of each study, as a complete replication is beyond the scope of this paper. Then we extend each analysis with the GC and HM approach ("quasi replication"; Bettis et al., 2016). The Stata code for estimating the GC approach as well as the code for our replication can be found in Online Appendices C and D, respectively.

Example 1: CSR and Shareholder Value in Mergers

Deng et al. (2013) used a sample of 1,556 mergers of U.S. firms from 1992 to 2007 to explore whether CSR creates value for acquiring firms in a merger context. They thereby juxtaposed two opposing theoretical views on CSR: the stakeholder value maximization view, which holds that CSR activities have a positive effect on shareholder wealth because focusing on the interests of other stakeholders increases their willingness to support a firm's operation, and the shareholder expense argument, which views CSR activities as an expense that is detrimental to shareholder value (Jahn & Brühl, 2018; Porter & Kramer, 2011).

Shareholder value is assessed via the cumulative abnormal shareholder returns of an event study, and the central explanatory variable, CSR activity, is measured using a composite CSR score based on indicators from the KLD Research & Analytics, Inc. STATS database. The study also incorporates various firm- and merger-specific control variables, as well as firm- and industry-fixed effects, to account for possible confounding effects.

Endogeneity concerns are important challenges for research of the CSR-FP relationship, and studies have questioned earlier findings that did not correct for endogeneity biases. For example, Zhao and Murrell (2016) reviewed the influential work by Waddock and Graves (1997) and showed that when correcting for endogeneity, financial performance still impacts subsequent CSR activities, but in contrast to the initial results by Waddock and Graves (1997), the effect of CSR on FP disappears. Similarly, Endrikat, Guenther, and Hoppe (2014) found that accounting for endogeneity changes the estimated relationship between CSR and FP.

Deng et al. (2013) explicitly addressed the endogeneity concerns of CSR research. They argued that the event study design mitigates the reverse causality problem as mergers are largely unanticipated events. However, despite an extensive list of control variables, omitted variable bias remains a concern: Better management teams (as an unobserved omitted variable) might be located in firms that undertake more profitable mergers and might also be investing more in CSR activities. It is also possible that if firms can predict merger opportunities, they might deliberately invest in CSR activities in anticipation of a merger. This would lead to problems of reverse causality as merger returns could reflect investments in CSR. To address these concerns, Deng et al. (2013) employed a

two-stage least squares regression, with one instrument being the religion ranking of the state where the acquirer's headquarters is located and the other being an indicator variable reflecting whether an acquirer's headquarters is based in a Democratic ("blue") state. Level of religiousness is positively correlated with attitudes toward CSR (Angelidis & Ibrahim, 2004), but this measure is unlikely to be related to a firm's merger performance. Similarly, firms with high CSR ratings tend to be located in Democratic or "blue" states (Rubin, 2008), but the decision to locate in a specific state should have no impact on merger performance. The authors' results support the stakeholder value maximization perspective, as they found that acquirers with a higher CSR realize higher merger announcement returns.

Example 2: International Competition and National Innovation

The relationship between competition and innovation has long been debated in management and economics (Gilbert, 2006). Although the standard economics perspective holds that higher competition leads to lower investment in innovations (Dasgupta & Stiglitz, 1980), recent studies suggest a complex and nuanced relationship between competition and innovation. For example, Aghion, Bloom, Blundell, Griffith, and Howitt (2005) showed an inverted U-shaped effect of competition on innovation once they accounted for firms' similarity, a previously omitted variable. In the context of firms facing increased international competition, Bloom, Romer, Terry, and Van Reenen (2014) showed that after an increase in Chinese import competition, European firms increased innovation. Autor et al. (2020) extended this debate and examined how greater exposure to trade from China impacts innovation of U.S. firms. The authors focused on the innovation activity (measured as utility patent) of firms in a U.S. manufacturing industry and how it can be explained by import exposure (measured by the Chinese import penetration ratio).

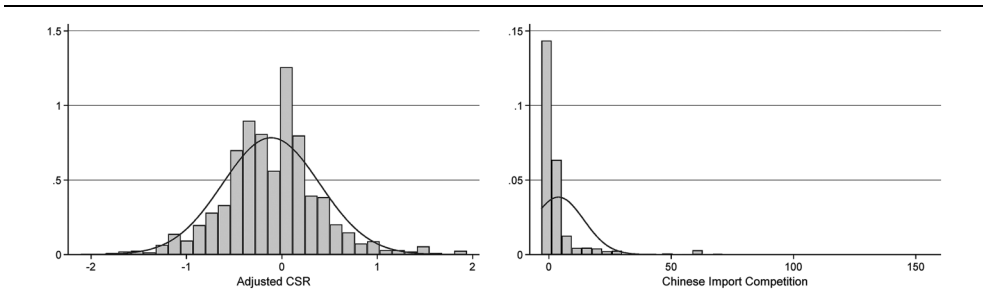
Autor et al. (2020) argued that endogeneity issues arise specifically due to concerns about the causality of the effects. In particular, domestic demand shocks in the United States could determine changes in import penetration and innovation. Furthermore, even if internal supply shocks are the main drivers of Chinese export growth, bilateral trade flow might still be influenced by U.S. industry import demand shocks. To account for this supply-driven component in U.S. imports from China, Autor et al. (2020) instrumented the Chinese import penetration in the United States via the import penetration in other high-income countries. The logic for this instrumental variable is that for high-income economies import demand shocks should largely be uncorrelated, whereas the exposure to imports growth from China is expected to be similar.

Overall, the results from Autor et al. (2020) indicate that industries or firms that are faced with a larger increase in trade exposure show smaller increases in patenting. Additionally, their analysis indicates that other key firm characteristics also decline with trade-exposed firms (i.e., global employment, sales, profitability, and R&D expenditure). This suggests that in the case of U.S. firms the primary response to import competition was to scale back global operations.

Replication and GC Comparison

As a first step, we examine the distribution of the endogenous regressors of the two studies against the normal distribution. Since our simulation studies in the previous section showed

Figure 2
Histograms of endogenous regressors in empirical examples



that relying on the Shapiro-Wilk test could lead to applications of the GC approach in cases where proper identification is not given, we propose that this test should always be used in combination with a visual inspection of the distribution of the endogenous regressor. Figure 2, therefore, shows a histogram with overlaid normal distributions for both endogenous regressors. The endogenous regressor appears to be highly nonnormal in Autor et al. (2020) and relatively normal in Deng et al. (2013). This is also supported by the associated skewness and kurtosis statistics. However, in both cases the results of the Shapiro-Wilk test reject the null hypotheses ($p < 0.00$), leading researchers to conclude that both distributions are not normal.

Table 6 Panel A shows the replication of the fully specified estimation results by Deng et al. (2013: Model 2 and 6 of Table 3). Similar to the original study, we present the baseline OLS (Model 1) followed by the second stage of IV regression (Model 2). We achieve an exact replication OLS Model and a very close replication of the IV results¹⁰ reported in Deng et al. (2013). In both estimations, the adjusted CSR score has a positive impact on the cumulative abnormal returns of the merger. However, in the GC estimation (Model 3) the adjusted CSR remains positive, but the impact is lower than in the IV regression and not significant. We also note that the additional GC regressor is not significant. Similarly, the HM estimation shows a positive and nonsignificant result; it does, however, lie between the IV and the GC estimate.

In Panel B we show the results for the replication of selective models of Autor et al. (2020). We thereby focus on the key periods of 1991–2007 (Autor et al., 2020: Table 2, column 3). Model 1 shows baseline OLS results without control variables and Model 3 the fully specified IV (Autor et al., 2020: Models A and H from Table 2, column 3). For better comparison with the other estimation methods, we also run an OLS estimation with control variables (Model 2), which was not part of the original study. Our estimation (Model 1 and Model 3) leads to an exact replication of the original study results (Autor et al., 2020: Models A and H from Table 2, column 3). Model 4 depicts the estimation of the GC estimation. In this case, the coefficient of the independent variables remains negative and significant with a level close to the OLS estimation in Model 2. Again, the additional GC regressor is not significant. The HM yields negative and insignificant results somewhat more distant from the prior OLS and IV estimations.

Table 6
Empirical Examples

	Panel A				Panel B				
	Deng et al. (2013)				Autor et al. (2020)				
Study	<i>Cum. abnormal return (-1,1) acquirer</i>				<i>Relative change in U.S. patents</i>				
Dependent Variable	<i>Adjusted CSR</i>				<i>Change of Chinese import penetration</i>				
Independent Variable	OLS	IV	GC	HM	OLS	OLS	IV	GC	HM
Estimation Model	1	2	3	4	1	2	3	4	5
Independent variable	0.500**	5.206**	2.132	2.926	-0.324	-1.070***	-1.353***	-0.980**	-0.673
	(0.249)	(2.370)	(1.854)	(7.685)	(0.258)	(0.174)	(0.497)	(0.449)	(0.645)
Gaussian Copula term			-0.866					-5.437	
			(1.037)					(9.053)	
Control variables	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	1,206	1,206	1,206	1,206	8,271	8,271	8,271	8,271	8,271
Hausman test (<i>p</i> value)		0.026					0.572		
Anderson LM test (<i>p</i> value)		0.000					0.034		
Hansen J test (<i>p</i> value)		1.000					0.000		

Note: Robust standard errors in parentheses; significance levels * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$. All models include full set of control variables as indicated by the original studies. In the case of the IV regression only the second stage is shown. Estimation of GC and HM estimation with bootstrapped standard errors. Panel A, Model 1 replicates Deng et al. (2013), Table 3, Model 2. Panel A, Model 2 replicates Deng et al. (2013), Table 3, Model 6. Standard errors are clustered by acquirer firm. Panel B, Model 1 replicates Autor et al. (2020), Table 2, column 3, Model A. Panel B, Model 1 replicates Autor et al. (2020), Table 2, column 3, Model H. We additionally estimate an OLS model with a full set of control variables (Panel B, Model 2). All models in Panel B are weighted by a firm’s U.S.-inventor patents averaged over the start and end of a period. Standard errors are clustered on four-digit SIC industries. The full Stata code for all estimations is available in Online Appendix D.

Overall, the results show that applying the GC approach without careful inspection of the regressor’s distribution can lead to incorrect inferences of nonsignificance. For the replication of Deng et al. (2013), where the regressor is close to normal, the GC estimator is insignificant and contradicts the results of the IV regression that uses strong and exogenous instruments as indicated by the Anderson LM test of underidentification and the Hansen J statistic of overidentification. The nonsignificance of the additional copula term—which contradicts the findings of the Hausman test—also reflects the findings from the simulations, which showed that nonsignificance of this term does not imply non-existing endogeneity.

Table 7
Summary of Gaussian Copula Approach

Key Assumptions

1. Enough variation in regressor
2. Nonnormality of regressor
3. Normality of error term
4. Dependence structure of the endogenous regressor and the error term can be described by a Gaussian copula

Advantages

- No need for instruments
- Extension to multiple endogenous regressors, as well as interactions possible
- Easy to implement
- Power similar to exogenous and strong IV regression if assumptions are met
- Confidence intervals include true values even if Assumption 2 or 3 is violated

Disadvantage

- Not applicable for sample selection and limited applicability for self-selection
- Impossible to test and difficult to justify Assumptions 3 and 4
- Strongly biased if Assumption 4 is not met

Recommendations Implementation and Reporting

- Do not use for sample selection or self-selection with limited number of categories
- Avoid use with low number of observations
- Discuss normality of endogenous regressor using statistical tests (e.g., Shapiro Wilk test or similar) and graphical assessments
- Highlight the limitation that the assumptions regarding the error term distribution and the error-regressor dependency structure are not testable
- Significance of additional GC regressor should not be used to confirm or reject presence of endogeneity
- Do not use as standalone method

In contrast, the replication of Autor et al. (2020) demonstrates the usefulness of the GC approach when the regressor is sufficiently nonnormally distributed. It shows that in these cases the GC approach can be useful as a robustness check for IV regressions: The instrument in Autor et al. (2020) is strong, as indicated by the Anderson LM test, but exogeneity cannot be tested as the number of instruments is equal to the number of endogenous regressors (Bascle, 2008). Additionally, the ability of the Hausman test to detect endogeneity is in question because this test depends on assuming exogenous instruments, among other things. In such a case, applying the GC approach can help to provide further confidence in the results, here supporting the OLS estimates.

Conclusions and Recommendations

The starting point of our study was the ongoing struggle of applied management research in dealing with endogeneity. Very frequently the use of IV methods is difficult or not feasible due to the lack of suitable instruments, often resulting in the use of instruments that lead to more harm than good (Murray, 2006; Semadeni et al., 2014). Management researchers thus increasingly turn to instrument-free methods, such as the GC approach, to deal with endogeneity (e.g., Becerra & Markarian, 2021; Bhattacharya et al., 2020; Reck et al., 2021). In this paper, we set out to evaluate the GC approach as an instrument-free method to address endogeneity for applied management researchers.

Throughout the paper, we iterate that the performance of the GC approach hinges on how well its assumptions are met and that a thorough understanding of these assumptions as well as their implications is a necessary condition for an appropriate application. Our first step was therefore to provide a thorough discussion of the GC assumptions in the context of management research. We then focused our simulations on three of the assumptions that are most likely to be violated in a management context—namely, the normal distribution of the error term, the nonnormal distribution of the endogenous regressor, and the dependency structure between error term and endogenous regressor. Our extensive set of simulation conditions provides several insights and recommendations for empirical management scholars dealing with endogeneity who hope to apply the GC approach in their research.

We find that each approach we considered has its specific advantages and disadvantages, and thus, we caution empirical research to consider carefully the trade-off of each. OLS clearly performs best in situations without endogeneity, but even low levels of endogeneity can lead to biased estimates. Similarly, IV approaches are very good if exogenous and strong instruments are available—yet such instruments are often hard to find or their quality difficult to prove. In contrast, the GC approach can provide estimates that are very close to the true values, but its performance relies on how well its assumptions are met. Although the GC approach may fail to detect significant relationships between the endogenous regressor and the dependent variable if the distribution of the regressor is not sufficiently nonnormal, the distribution of the regressor can be assessed by researchers and thus precautions can be taken. More concerning, however, is the fact that the GC approach may lead to estimates that are more biased than those obtained with exogenous, weak instruments when the error term is skewed. Additionally, the GC approach leads to biased estimates that can be farther away from the true parameter than those obtained with regressions with endogenous instruments when the data cannot be described by a Gaussian copula. We argue that both of these untestable assumptions can hold in a variety of situations, but we also caution researchers not to blindly accept them.

Connecting our study to prior research and gearing it to applied researchers, we summarize our findings in Table 7 and suggest the following procedure when implementing the GC approach (see also decision tree in Online Appendix E). First, researchers must specify a theoretically grounded econometric model and ensure that they include a comprehensive set of control variables in the regression. They should also consider transforming variables if they suspect the error terms to be nonnormal (see Schmidt & Finan, 2018, for a more detailed discussion). Previous research suggested using the distribution of the residuals as an *indicator* for the shape of the error term distribution (Becker et al., 2021; Papies et al., 2016), yet we caution researchers against doing so as residuals are a function of model (mis-)specification (McKean et al., 1993) and thus of arguably less use in the presence of endogeneity. Researchers should then carefully assess whether endogeneity could be a problem. If the only purpose of the model is predictive (i.e., to provide forecasts), there is no need to correct for endogeneity at all (Rossi, 2014). In situations where the estimates are of primary interest, however, the researcher faces a challenge as the error term is unobservable, which implies that there is no possible way to directly test whether it is correlated with the regressor (Ullah, Akhtar, & Zaefarian, 2018). Statistics such as the Impact Threshold of a Confounding Variable or the Robustness of Inference to Replacement may help researchers determine whether there is sufficient concern that endogeneity bias may overturn the insights

obtained from a (misspecified) regular OLS (Busenbark, Frank, Maroulis, Xu, & Lin, 2021; Busenbark et al., 2022). Yet since interpretations of these statistics rely to a certain degree on subjective judgments, we recommend always checking the OLS results against methods that account for endogeneity.

The next step for any researcher suspecting that endogeneity is present is therefore to determine where this endogeneity stems from, since the causes of endogeneity determine the suitability of the different remedies to address it. However, the general review of endogeneity by Hill et al. (2021) and our reading of the existing studies in management using the GC approach both indicate that a substantial number of studies do not specify the source of endogeneity precisely. This is particularly concerning as the GC approach should not be used for sample-selection-based endogeneity and is not suitable for most forms of self-selection bias as discussed in Hamilton and Nickerson (2003) and Clougherty, Duso, and Muck (2016), among others. Although the GC approach can in theory be applied in these situations, the scant information included in the different categories of the endogenous regressor prevents the efficient use of the CG approach.

If neither sample-selection-based nor self-selection-based endogeneity is suspected, the availability and quality of suitable instruments and distribution of the endogenous regressor are crucial for deciding between the IV approach and instrument-free methods. The fact that strong instruments are hard to find and their exogeneity often difficult or even impossible to test (Rossi, 2014), as well as the assumptions of most instrument-free methods being challenging or impossible to verify (Lewbel, 2019), leads us to recommend a prudent approach when attempting to control for endogeneity and validate the IV regression results with those of the chosen instrument-free method and vice versa.

A possibly suitable instrument-free method if the regressor is nonnormal is the GC approach. Although nonnormality of the regressor can be easily observed, we also note that it limits the application of the GC approach as many important potential regressors are close to normally distributed; examples include stock market returns measured as cumulative abnormal return within event study frameworks (Sorescu et al., 2017) and psychological variables like the intelligence quotient (Mendelson, 2000). Yet many organizational phenomena rely on variables that are not normally distributed: use of alliances, R&D spending, patent quality, and firm size, just to name a few (e.g., Almeida et al., 2011; Hohberger, 2016; Segarra & Teruel, 2012). Furthermore, in some instances normally distributed measures may be replaced with nonnormal measures even when examining the same research question. For example, our empirical example reproducing Deng et al. (2013) relies on the aggregated KLD scores, which are relatively normally distributed and thus violate a key assumption of the GC approach. However, the underlying categories of the KLD scores are nonnormally distributed (Erhemjamts, Li, & Venkateswaran, 2013) and they can also be used as predictors (Chava, 2014). In any case, there should still be ample potential for applications of the GC approach in management research despite its distributional requirement. We note that it is also possible to combine the GC approach with IV by using instruments to account for the set of endogenous regressors for which instruments are available and using GC controls for the remaining endogenous regressors, thus benefiting from the advantages of both approaches.

Reporting statistical procedure and data suitability is an important part of the correct application of the statistical approaches to control for endogeneity. In the context of IV approaches, Semadeni et al. (2014) recommended that authors should always report the strength (using the

F-statistic from the first stage) and the exogeneity (using, for example, the Sargan test, 1958) of chosen instruments. Similarly, we suggest providing a brief description of the shape of the endogenous regressor and highlighting the assumptions regarding the error term. This is particularly important as our simulations show that inappropriately applying the GC approach may result in parameters that are biased or for which significance tests can be misleading. To test the shape of the regressor, it is advisable to apply the Shapiro-Wilk test or similar tests for non-normality. Ideally, researchers should also complement these test statistics with a visual assessment of the regressor's distribution.

The assumptions of the GC approach regarding the distribution of the error term as well as the dependency structure between endogenous regressor and error term cannot be tested. Given their importance for the performance of the GC approach, we nevertheless strongly encourage researchers to highlight them to ensure a balanced reporting of their results.

Finally, we also suggest reporting the significance of the GC correction regressor. However, while recent management applications used the GC correction term to confirm (e.g., Becerra & Markarian, 2021) or reject endogeneity (e.g., Reck et al., 2021), our simulations suggest a more conservative interpretation. We show that the significance of the GC correction term may be indicative of endogeneity when the GC assumptions are met. Yet it should be interpreted with caution as—particularly when the assumptions of the GC approach are not met—this additional regressor does not show up as significant despite endogeneity being present. Similarly, in a small number of cases, the GC correction regressor is significant despite the absence of endogeneity. As such, relying on the GC correction regressor to rule out or establish endogeneity may provide false confidence in estimation results.

In summary, our paper provides important insights into the usefulness of the GC approach. In particular, in cases where instruments are scarce, or their suitability is difficult to assess, the GC approach can be a useful and practical alternative. Yet we caution researchers against its unreflected use and encourage them to carefully weigh the disadvantages of this approach. Like other instrument-free approaches (Lewbel, 2019), the GC approach should ideally be employed in conjunction with other means of obtaining identification, both as a way of checking the robustness of results and of increasing the efficiency of estimation. Overall, we believe the GC approach should be seen as an addition to the existing set of statistical tools to address endogeneity. However, as with any other method, it has its limitations and is not the cure for all endogeneity concerns.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jan Hohberger  <https://orcid.org/0000-0002-1177-8147>

Notes

1. Becker et al. (2021) derived a decision tree that provides suggestions on the usefulness of the GC approach given sample size and regressor skewness. While such a decision tree is certainly helpful, we note that the goal of the decision tree is to predict an 80% power of the copula correction term ($\hat{\alpha}$) and not a loss function that captures the bias of the estimates.
2. Table 1 is partly based on the discussion in Ebbes et al. (2009). It provides a brief summary of the different methods, as does the discussion that follows next, but note that it is outside the scope of this paper to provide a full picture of these methods. Interested readers should thus refer to the original research papers.
3. Note that two of the six suggested constructed instruments in Lewbel (1997) require the model error and the measurement error, respectively, to be symmetrically distributed.
4. In our simulations we dynamically adjust the correlation in Step 1 if the actual correlation between x and e deviates more than 5% from the desired correlation.
5. For the implementation of the HM approach, we used all three instruments discussed in Ebbes et al. (2009) and obtained standard errors via bootstrapping. For all estimation methods we included an intercept in the estimation to provide a true picture of the GC's performance (see also Becker et al., 2021).
6. The HM approach requires the regressor to be asymmetric. Yet its performance in the no-skew condition is better than expected insofar as the t-bias statistic shows that the estimates—though far away the true value—are still less than 2 SE away from the true values.
7. Becker et al. (2021) suggested using Anderson-Darling and Cramer-van Mises normality tests, though also note that these may not be fully able to rule out nonidentification.
8. Although in case of nonnormal error terms the regular t-statistic may not be useful (Pek, Wong, & Wong, 2018), we continue to interpret the t-bias statistic as before.
9. We would like to thank the authors of both studies for providing their data for replication and Dr. Xin Deng for providing additional insights in our replication of Deng et al. (2013).
10. In contrast to Deng et al. (2013), we use a full set of control variables in the first stage of the IV regression.

References

- Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. 2005. Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics*, 120: 701-728.
- Almeida, P., Hohberger, J., & Parada, P. 2011. Individual scientific collaborations and firm-level innovation. *Industrial and Corporate Change*, 20: 1571-1599.
- Andries, P., & Hünermund, P. 2020. Firm-level effects of staged investments in innovation: The moderating role of resource availability. *Research Policy*, 49: 103994.
- Angelidis, J., & Ibrahim, N. 2004. An exploratory study of the impact of degree of religiousness upon an individual's corporate social responsiveness orientation. *Journal of Business Ethics*, 51: 119-128.
- Autor, D., Dorn, D., Hanson, G. H., Pisano, G., & Shu, P. 2020. Foreign competition and domestic innovation: Evidence from US patents. *American Economic Review: Insights*, 2: 357-374.
- Bascle, G. 2008. Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organization*, 6: 285-327.
- Becerra, M., & Markarian, G. 2021. Why are firms with lower performance more volatile and unpredictable? A vulnerability explanation of the bowman paradox. *Organization Science*, 32: 1149-1390.
- Becker, J.-M., Proksch, D., & Ringle, C. M. 2021. Revisiting Gaussian copulas to handle endogenous regressors. *Journal of the Academy of Marketing Science*, 50: 46-66.
- Bettis, R. A., Helfat, C. E., & Shaver, J. M. 2016. The necessity, logic, and forms of replication. *Strategic Management Journal*, 37: 2193-2203.
- Bhattacharya, A., Good, V., Sardashti, H., & Peloza, J. 2020. Beyond warm glow: The risk-mitigating effect of corporate social responsibility (CSR). *Journal of Business Ethics*, 171: 1-20.
- Bitrián, P., Buil, I., & Catalán, S. 2021. Enhancing user engagement: The role of gamification in mobile apps. *Journal of Business Research*, 132: 170-185.
- Bloom, N., Romer, P., Terry, S. J., & Van Reenen, J. 2014. Trapped factors and China's impact on global growth. *The Economic Journal*, 131: 156-191.

- Busenbark, J. R., Frank, K. A., Maroulis, S. J., Xu, R., & Lin, Q. 2021. Quantifying the robustness of empirical inferences in strategic management: The impact threshold of a confounding variable and robustness of inference to replacement. In H. Aaron, M. Aaron, O. K. Paula, & S. Paroutis (Eds.), *Research Methodology in Strategy and Management: Research in Crisis*, Vol. 13: 123-150. Bingley, UK: Emerald Publishing.
- Busenbark, J. R., Yoon, H., Gamache, D. L., & Withers, M. C. 2022. Omitted variable bias: Examining management research with the impact threshold of a confounding variable (ITCV). *Journal of Management*, 48: 17-48.
- Certo, S. T., Busenbark, J. R., Woo, H. S., & Semadeni, M. 2016. Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal*, 37: 2639-2657.
- Certo, S. T., Raney, K., & Albader, L. 2020. The non-normality of firm performance and the consequences for strategic management theory. *Academy of Management Proceedings*, 1: <https://doi.org/10.5465/AMBPP.2020.5291>.
- Chava, S. 2014. Environmental externalities and cost of capital. *Management Science*, 60: 2223-2247.
- 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: 286-347.
- Cohen, J. 1988. *Statistical Power Analysis for the Social Sciences*. Hillsdale, NJ: Erlbaum.
- Cortina, J. M., & Dunlap, W. P. 1997. On the logic and purpose of significance testing. *Psychological Methods*, 2, 161-172.
- Danaher, P. J., & Smith, M. S. 2011. Modeling multivariate distributions using copulas: Applications in marketing. *Marketing Science*, 30: 4-21.
- Dasgupta, P., & Stiglitz, J. 1980. Industrial structure and the nature of innovative activity. *The Economic Journal*, 90: 266-293.
- Datta, H., Ailawadi, K. L., & van Heerde, H. J. 2017. How well does consumer-based brand equity align with sales-based brand equity and marketing-mix response? *Journal of Marketing*, 81: 1-20.
- Deng, X., & Kang, J.-k., & Low, B. S. 2013. Corporate social responsibility and stakeholder value maximization: evidence from mergers. *Journal of Financial Economics*, 110: 87-109.
- Durante, F., Fernandez-Sanchez, J., & Sempì, C. 2013. A topological proof of Sklar's theorem. *Applied Mathematic Letters*, 26: 945-948.
- Ebbes, P., Wedel, M., & Böckenholt, U. 2009. Frugal IV alternatives to identify the parameter for an endogenous regressor. *Journal of Applied Econometrics*, 24: 446-468.
- Ebbes, P., Wedel, M., Böckenholt, U., & Steerneman, T. 2005. Solving and testing for regressor-error (in) dependence when no instrumental variables are available: With new evidence for the effect of education on income. *Quantitative Marketing and Economics*, 3: 365-392.
- Endrikat, J., Guenther, E., & Hoppe, H. 2014. Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance. *European Management Journal*, 32: 735-751.
- Erickson, T., & Whited, T.M.. 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. *Econometric Theory*, 18: 776-799.
- Erhemjants, O., Li, Q., & Venkateswaran, A. 2013. Corporate social responsibility and its impact on firms' investment policy, organizational structure, and performance. *Journal of Business Ethics*, 118: 395-412.
- Falkenström, F., Park, S., & McIntosh, C. N. 2021. Using copulas to enable causal inference from non-experimental data: tutorial and simulation studies. *Psychological Methods*, (Forthcoming).
- Gilbert, R. 2006. Looking for Mr. Schumpeter: Where are we in the competition-innovation debate? *Innovation Policy and the Economy*, 6: 159-215.
- Guitart, I. A., & Stremersch, S. 2021. The impact of informational and emotional television ad content on online search and sales. *Journal of Marketing Research*, 58: 299-320.
- Hamilton, B. H., & Nickerson, J. A. 2003. Correcting for endogeneity in strategic management research. *Strategic Organization*, 1: 51-78.
- Hasan, M. M., Taylor, G., & Richardson, G. 2021. Brand capital and stock price crash risk. *Management Science*, 1-27.
- Haschka, R. E., & Herwartz, H. 2020. Innovation efficiency in European high-tech industries: Evidence from a Bayesian stochastic frontier approach. *Research Policy*, 49: 104054.
- Heckman, J. J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica: Journal of the Econometric Society*, 4: 931-959.

- Hill, A. D., Johnson, S. G., Greco, L. M., O'Boyle, E. H., & Walter, S. L. 2021. Endogeneity: A review and agenda for the methodology-practice divide affecting micro and macro research. *Journal of Management*, 47: 105-143.
- Hohberger, J. 2016. Does it pay to stand on the shoulders of giants? An analysis of the inventions of star inventors in the biotechnology sector. *Research Policy*, 45: 682-698.
- Huang, W., & Prokhorov, A. 2014. A goodness-of-fit test for copulas. *Econometric Reviews*, 33: 751-771.
- Jahn, J., & Brühl, R. 2018. How Friedman's view on individual freedom relates to stakeholder theory and social contract theory. *Journal of Business Ethics*, 153: 41-52.
- Larcker, D. F., & Rusticus, T. O. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, 49: 186-205.
- Lewbel, A. 1997. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D. *Econometrica: Journal of the Econometric Society*, 65: 1201-1213.
- Lewbel, A. 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30: 67-80.
- Lewbel, A. 2019. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57: 835-903.
- McKean, J. W., Sheather, S. J., & Hettmansperger, T. P. 1993. The use and interpretation of residuals based on robust estimation. *Journal of the American Statistical Association*, 88: 1254-1263.
- Mendelson, H. 2000. Organizational architecture and success in the information technology industry. *Management Science*, 46: 513-529.
- Murray, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *The Journal of Economic Perspectives*, 20: 111-132.
- Nelsen, R. B. 2007. *An introduction to Copulas*. New York, NY: Springer Science & Business Media.
- Papies, D., Ebbes, P., & Van Heerde, H. J. 2016. Addressing endogeneity in marketing models. In P. S. H. Leeflang, J. E. Wieringa, T. H. A. Bijmolt, & K. H. Pauwels (Eds.), *Advanced Methods for Modeling Markets* : 581-627. Switzerland: Springer.
- Park, S., & Gupta, S. 2012. Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, 31: 567-586.
- Pek, J., Wong, O., & Wong, A. 2018. How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology*, 9: 2104.
- Porter, M. E., & Kramer, M. R. 2011. Creating shared value. *Harvard Business Review*, 89: 62-77.
- Reck, F., Fliaster, A., & Kolloch, M. 2021. How to build a network that facilitates firm-level innovation: An integration of structural and managerial perspectives. *Journal of Management Studies*, 1-34.
- Rigobon, R. 2003. Identification through heteroskedasticity. *Review of Economics and Statistics*, 85: 777-792.
- Rossi, P. E. 2014. Even the rich can make themselves poor: A critical examination of IV methods in marketing applications. *Marketing Science*, 33: 655-672.
- Rubin, A. 2008. Political views and corporate decision making: The case of corporate social responsibility. *Financial Review*, 43: 337-360.
- Ruscio, J., & Kacetow, W. 2008. Simulating multivariate nonnormal data using an iterative algorithm. *Multivariate Behavioral Research*, 43: 355-381.
- Rutz, O. J., & Watson, G. F. 2019. Endogeneity and marketing strategy research: An overview. *Journal of the Academy of Marketing Science*, 47: 479-498.
- Schmidt, A. F., & Finan, C. 2018. Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, 98: 146-151.
- Segarra, A., & Teruel, M. 2012. An appraisal of firm size distribution: Does sample size matter? *Journal of Economic Behavior & Organization*, 82: 314-328.
- 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: 1070-1079.
- Shapiro, S. S., & Wilk, M. B. 1965. An analysis of variance test for normality. *Biometrika*, 52: 591-611.
- Shaver, J. M. 2020. Causal identification through a cumulative body of research in the study of strategy and organizations. *Journal of Management*, 46: 1244-1256.
- Sklar, M. 1959. Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8: 229-231.

- Sonnier, G. P., McAlister, L., & Rutz, O. J. 2011. A dynamic model of the effect of online communications on firm sales. *Marketing Science*, 30: 702-716.
- Sorescu, A., Warren, N. L., & Ertekin, L. 2017. Event study methodology in the marketing literature: An overview. *Journal of the Academy of Marketing Science*, 45: 186-207.
- Ullah, S., Akhtar, P., & Zaefarian, G. 2018. Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*, 71: 69-78.
- Waddock, S. A., & Graves, S. B. 1997. Quality of management and quality of stakeholder relations: Are they synonymous? *Business & Society*, 36: 250-279.
- Wooldridge, J. M. 2012. *Introductory Econometrics: A Modern Approach*. Toronto: Nelson Education.
- Zhao, X., & Murrell, A. J. 2016. Revisiting the corporate social performance-financial performance link: A replication of waddock and graves. *Strategic Management Journal*, 37: 2378-2388.n