MARKETING
ZFP – Journal of Research and Management
www.marketing-zfp.de

Research Articles

Control Variables in Marketing Research

By Martin Klarmann and Sven Feurer

In empirical marketing research that does not rely on fully randomized experiments, control variables are an important tool to rule out rival alternative explanations for the observed relationships. Despite their importance for causal inference, control variables often receive little attention from either applied researchers or methodologists. At the same time, overviews of control variable practices in neighboring disciplines demonstrate that researchers struggle with selecting, analyzing, and interpreting control variable results. In response, this article combines a synthesis of the theoretical knowledge on control variables with a review of control variable practices. Against this background, we develop and discuss sixteen recommendations for control variable use in marketing research.

1. Introduction

In recent years, marketing researchers have become progressively more interested in whether their empirical results can be interpreted causally. For instance, "Quantitative models to understand causality, levers, and influences in a complex world" is one of the Marketing Science Institute's 2016–2018 top research priorities (Marketing Science Institute 2016). Marketing is not alone in this quest: "In many scientific fields such as economics, psychology, education, and environmental science, statistical models are used almost exclusively for causal explanation" (Shmueli 2010, p. 289). In this vein, major textbooks on causal inference have been published recently, covering the social sciences in general (Morgan and Winship 2015), economics (Angrist and Prischke 2015; Imbens and Rubin 2015), and epidemiology (VanderWeele 2015). For marketing, causal inference is particularly important, because researchers are asked to come up with "actionable implications that would capture the attention of the practitioner community" (Kumar 2016, p. 6). The causal statement "Action A leads to outcome B" is more actionable than the correlational statement "Action A is positively related to outcome B."

The benchmark for causal inference is the randomized experiment (Koschate-Fischer and Schandelmeier 2014). However, in many important research fields in marketing – including marketing strategy, B2B marketing, sales, and marketing-mix modeling – conducting experiments is typically not feasible. Consider, for instance, organizational downsizing (e. g., Habel and Klarmann 2015; Homburg et al. 2012b). Researchers obviously cannot randomly assign firms to conditions such that firms in one condition must lay off employees, whereas firms in the other condition must not. Hence, researchers in these fields turn to next best alternatives, such as examining survey data, archival data, data from social media, and data from quasi-experiments.

In all these settings, control variables are an important tool for researchers to meet one criterion for causal inference (of many): the exclusion of alternative explanations for the observed relationship. The idea is that by adding control variables to the model, hypothesized effects are estimated at constant levels of the control variables. If hypothesized relationships still hold after adding the controls, alternative causal explanations involving the control variables can be ruled out. Reversely, if alternative explanations are not accounted for in the model, the analysis suffers from "omitted variable bias" or "endogeneity" (Ebbes et al. 2017).

Hence, theoretically, the case for including control variables to improve causal interpretability is straightforward. However, from a practical perspective little seems to be clear. As Carlson and Wu (2012, p. 414) note, issues surrounding control variables "may not be widely


Martin Klarmann is Professor of Marketing at the Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Zirkel 2, 76049 Karlsruhe, Germany. Phone: +49 721 608 43726, E-mail: martin.klarmann@kit.edu. *Corresponding Author.

Sven Feurer is Postdoctoral Researcher at the Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Zirkel 2, 76049 Karlsruhe, Germany. Phone: +49 721 608 43726, E-mail: sven.feurer@kit.edu.

Please note: The authors thank Kelly L. Haws, Dominik Papies, and members of the review team for their helpful advice.

understood or are systematically ignored by authors and reviewers." A number of recent studies (mostly in organizational behavior) seeking to improve researchers' control variable use list several practical uncertainties (e. g., Atinc et al. 2012; Bernerth and Aguinis 2016; Spector and Brannick 2011). Additionally, in their much-cited study on false positive psychology, Simmons et al. (2011) identify control variable selection as an aspect of „researcher degrees of freedom." As which control variables need to be included is often unclear, researchers may play around with different sets of control variables until they receive the results they like – a practice sometimes referred to as "p-hacking" (Simonsohn et al. 2014, p. 670).

Against this backdrop, our goal is to develop recommendations for good control variable use in marketing research. Instead of relegating control variables to a short paragraph – the usual treatment in empirical studies and textbooks – we make control variables the focus of our paper. The paper's primary contribution lies in combining formal analysis with a discussion of practical issues surrounding control variable use. A formalized and normative discussion of control variables can often be found in textbooks on econometrics, but practical concerns are rarely addressed. At the same time, review studies of control variable use in empirical research practice typically take a more descriptive empirical approach.

The principal field of application of our paper is cross-sectional survey research, where claims about causality are most difficult (Bono and McNamara 2011; Sande and Ghosh 2018). However, many of our recommendations generalize to other forms of data, some even to randomized experiments. Moreover, while our paper analyzes the problem from the perspective of OLS regression, many findings generalize to other multivariate methods, particularly analysis of (co-)variance and maximum-likelihood estimation of structural equation models.

To achieve our goal, we begin by adopting a theoretical stance. Drawing on econometrics and causal graph analysis, we analyze what control variables can and cannot accomplish. Subsequently, we review the relevant literature on current control variable practice in the social sciences. This review guides us in identifying applied issues in control variable use. Finally, we derive recommendations for good control variable use in marketing research.

2. Control variable theory

From a theoretical perspective, two motivations support adding variables to an empirical model that are not directly evoked by the hypotheses of the researchers. First and foremost, including control variables allows for an improved causal interpretability of the estimated coefficients. Theorists on causal inference typically list several criteria for a causal interpretation of observed relationships. Drawing on a synthesis of the literature, Edwards and Bagozzi (2000) define four: (1) cause and effect must be distinct phenomena, (2) cause and effect must be associated, (3) the cause must precede the effect in time, and (4) alternative causal explanations of the observed relationship must be ruled out. Control variables are used to meet the fourth criterion. They capture rival causal explanations for a study's focal relationships.

As an illustrative application of control variables, consider the study of Germann et al. (2015). The authors are interested in whether firms with a chief marketing officer (CMO) outperform firms without this officer. However, having a strong corporate brand quite possibly increases both the probability of having a CMO and the firm's performance. A strong brand could therefore imply a positive correlation between CMO presence and performance, without any causal link. As a response, Germann et al. (2015) include "corporate branding" as a control variable (along with many others). They still find a significant effect of CMO presence on firm performance, and conclude that "the chief marketing officer matters!"

While this paper is mostly interested in how this first motivation extends to improving the interpretability of main effects of the independent variable x on a dependent variable y, several related applications are worth noting. Importantly, a well-established practice in applied marketing research is the inclusion of the simple effect of a moderator variable in a model that involves an interaction between moderator and independent variable. Otherwise, the effect of the interaction term may wrongfully capture the effect of the component that is omitted (e. g., Irwin and McClelland 2001). Moreover, the general logic behind the use of control variables also extends to variables that are included in a model to justify the missing-at-random assumption some approaches require for dealing with missing data (Thoemmes and Rose 2014).

Of note is that this causal logic extends only to research that aims to explain (vs. predict) why certain phenomena are linked (Shmueli 2010). In predictive applications (such as many machine learning tasks; Hofman et al. 2017), this reasoning does not apply.

The second motivation for using control variables is to improve the precision of the estimated coefficients. The additional variables in the model are thought to be important explanatory variables for the dependent variable while being uncorrelated with the hypothesized independent variables. Their inclusion in the model explains statistical noise in the dependent variable, increasing the precision of the estimated coefficients of the hypothesized effects.

In applied marketing research, these additional variables are typically referred to as "control variables," "confounding variables," and "covariates." Sometimes they are also referred to as "nuisance variables" (Kirk 1995), „disturbers" (Steyer and Schmitt 1994), and "concomitant variables" (Pratt and Schlaifer 1988). Strictly speaking, the term "control variables" seems appropriate only for variables added for the first reason. However, applied research is flexible in this regard, and so are we.

In the following two subsections, we formally analyze these two motivations that are driving control variable use in marketing research. As marketing researchers put more emphasis on the causal meaning of their claims than on the precision of estimated effect sizes, the first subsection is more detailed.

2.1. Control variables to improve the causal interpretability of results

2.1.1. Econometric analysis of omitted variable bias

A good way of looking formally at this motivation behind control variable use is to analyze how well regression coefficients recover true causal relationships. We follow the approach that is presented in Cameron (2006). We assume that a researcher is interested in the causal effect β of an independent variable x on the dependent variable y. The true causal model (often also referred to as „data-generating process") linking x to y is

$$y = \alpha + \beta x + \varepsilon \tag{1}$$

where α is the expected value for y if x is zero and ε are the residuals that capture the variation in y that x does not explain. The bivariate regression model that is actually analyzed is

$$y = a + bx + e \text{ with the OLS estimate} \tag{2}$$

$$b = \frac{Cov(x,y)}{Var(x)} \tag{3}$$

Notably, ε in (1) is a disturbance term, and e in equation (2) is a residual term. Conceptually, both are quite different. In particular, by construction, e is uncorrelated with x after OLS estimation. Hence, it becomes important to see to what extent b recovers the true effect β. For that purpose, we can insert (1) into (3), leading to:

$$b = \frac{Cov(x,\alpha + \beta x + \varepsilon)}{Var(x)} = \frac{Cov(x,\alpha) + Cov(x,\beta x) + Cov(x,\varepsilon)}{Var(x)} \tag{4}$$

As α is a constant, Cov(x,α) = 0. Moreover, Cov(x,βx) = βVar(x). Hence, (4) becomes:

$$b = \beta + \frac{Cov(x,\varepsilon)}{Var(x)} \tag{5}$$

Hence, β recovers the true effect of x on y only if Cov(x,ε) = 0. This is the well-known exogeneity assumption of OLS regression (e. g., Ebbes et al. 2017). Importantly, it is practically impossible to test whether this assumption is violated without making other untestable assumptions about the data-generating process.

Omitting variables that are linked to both x and y („common causes") from the regression creates a situation in which Cov(x,ε) ≠ 0. It needs to be emphasized that omitted variables are not the only reason for endogeneity. Ebbes et al. (2017) discuss other potential reasons.

To better understand omitted variable bias (and the use of control variables), we analyze how omitting a common cause of y and z can create a non-zero covariance between x and ε (we are following Bollen 1989, pp. 45–56).

Let us assume that a third variable z is the cause of x. Moreover, in addition to x, z is also a cause of y. Hence, we have:

$$y = \alpha_y + \beta_1 x + \beta_2 z + \varepsilon_y \text{ and} \tag{6}$$

$$x = \alpha_x + \gamma z + \varepsilon_x \tag{7}$$

If we omit z from the model in (6), β_2z becomes part of the disturbance term. That is, $\varepsilon_{y,new} = \beta_2 z + \varepsilon_y$. If we assume that Cov(z, ε_y=0), then we also know from (7) that Cov(x,z) = γVar(z). Hence,

$$\begin{aligned} Cov(x,\varepsilon_{y,new}) &= Cov(x,\beta_2 z) + Cov(x,\varepsilon_y) \\ &= \beta_2 \gamma Var(z) + Cov(x,\varepsilon_y) \neq 0_i \\ &= \alpha_x + \gamma z_i + \varepsilon_{x,i} \end{aligned} \tag{8}$$

In other words, if we omit z from the regression and just estimate (2), b will not recover the true effect β_1. Instead, b will be biased:

$$b = \beta_1 + \beta_2 \gamma \frac{Var(z)}{Var(x)} \tag{9}$$

Equation (9) has a number of important implications for how omitted variables can create bias. In particular, b captures multiple effects, not only β_1. This effect can create situations in which b is substantial, even if $\beta_1 = 0$. At the same time, if $\beta_2 < 0$ and $\beta_1 > 0$ (or vice versa), b could be estimated as zero, even if x has a causal effect. Possibly, including a variable may even change the sign of b, a phenomenon that has received quite a bit of attention as "Simpson's paradox" (e. g., Pearl 2014). In short, we can no longer rely on b if we omit z.

As (9) shows, the bias will disappear under only two conditions. First, if z is unrelated to y, then $\beta_2 = 0$ and no bias is present. Omitting covariates of x that are unrelated to y will not bias the estimate for the effect of x on y. Intuitively, this is clear: A variable that does not affect y cannot be an alternative explanation for any observed relationship between x and y. This condition has implications for the control variable selection process: Only variables that can plausibly affect y should be included as controls.

Second, if z is unrelated to x, then $\gamma = 0$ and no bias is present. In other words, omitting variables that are unrelated to x will not bias the estimate for the effect of x on y. Again, this circumstance has implications for the control variable selection process. Only variables that can plausibly affect x should be included as controls.

2.1.2. Econometric motivation of control variable use

The purpose of using control variables is to remove the bias shown in (9) by including z in the model that is analyzed. When we add z to the analyzed model, the bias in b (then: b_1) disappears. The model that is analyzed in this case is:

$$y = a + b_1x + b_2z + e \text{ with the OLS estimate} \quad (10)$$

$$b_1 = \frac{s_y}{s_x}\frac{r_{yx} - r_{yz}r_{xz}}{1 - r_{xz}^2} \text{ (Eid et al. 2010, p. 608)} \quad (11)$$

To simplify the analysis, we look at the case where Var(x) = Var(z) = Var(y) = 1. The formulas are then less complex, but the implications of control variable use do not change.

We continue to assume Cov(z, ε_y=0), implying that $r_{xz}=\gamma$. By substituting y in (11) with (6) and by substituting r_{xz} with γ we get:

$$b_1 = \frac{\beta_1 + \beta_2\gamma + Cov(x,\varepsilon_y) - (\beta_2 + \beta_1\gamma + Cov(z,\varepsilon_y))\gamma}{1-\gamma^2}$$

$$= \beta_1 + \frac{Cov(x,\varepsilon_y)}{1-\gamma^2} - \frac{\gamma Cov(z,\varepsilon_y)}{1-\gamma^2} \quad (12)$$

Thus, by including z in the regression of y on x, b_1 is unbiased. This basic motivation underlies the use of control variables to improve the causal interpretability of parameters.

Three assumptions need to be addressed. To get $b_1=\beta_1$, the first assumption that is needed is ($Cov(x,\varepsilon_y) = 0$). Thus, when controlling for z, x needs to be exogenous to y. This requirement may be violated if there are other omitted variables or if other sources of endogeneity are present (e. g., reverse causality). The second assumption is that ($Cov(z,\varepsilon_y) = 0$). Hence, z needs to be exogenous to y so that b_1 estimates β_1. This assumption has more important implications for control variable selection. It means that common causes of x and y included in the model need to be "last" in the sense that no other phenomena cause z and y. Also, a reciprocal relationship must not exist between y and the common causes.

A third assumption that we made before deriving these results is that z is exogenous to x ($Cov(z,\varepsilon_x) = 0$). However, this assumption can be relaxed somewhat. If omitted common causes of z and x cause z and ε_x to covary, of sole importance is that the covariance between z and x captures all variation in x explained by the external cause that also causes the control variable. Thus, if a common cause explains both x and z, then $Cov(z,\varepsilon_x) \neq 0$ is implied. While γ will be biased as an estimate for the causal effect of z on x, b_1 will still be unbiased. (We return to this issue in subsection 2.1.6 when describing causal path analysis as a tool to identify suitable controls.)

However, if $Cov(z,\varepsilon_x) \neq 0$ owing to other forms of endogeneity, this is a problem. Of particular importance is that Cov(z,x) does not capture any causal influence of x on z. This issue has received some attention in the literature in the form of a distinction between pre-treatment control variables and post-treatment control variables (e. g., Angrist and Prischke 2015, pp. 214–217; Gelman and Hill 2007, pp. 188–190). The idea is that to aid causal inference, control variables need to be measured before x. Otherwise they could be affected by x. This measurement is particularly important for experiments, where – in principle – all controls should be measured before the treatment.

2.1.3. Consequences of measurement error in control variables

As will become evident in section 3, control variables used in applied empirical business research do not always meet the requirements we have laid out. Moreover, researchers often exercise less care in measuring control variables reliably than they do in measuring their focal variables. In this subsection, we are interested in the extent to which the presence of random measurement error in control variables constitutes a problem. Systematic measurement error in control variables is of lesser interest – not because the consequences are less problematic (they are not), but because the presence of systematic measurement error clearly results in an unpredictable bias such that any attempts to infer causality will fail (Homburg et al. 2012c).

For the following analysis in the tradition of the classical true score model, we assume that z (but not x) is measured with error. The degree to which z is free of measurement error is typically referred to as reliability $\rho_{zz'}$, which is defined as (e. g., McDonald 1999, p. 65):

$$\rho_{zz'} = \frac{Var(\xi)}{Var(z)} \quad (13)$$

Now, z is the observed variable that measures the true common cause of x and y: ξ. In particular, z measures ξ with random error δ that we assume has a mean of zero and is random in the sense that it is not linked to any phenomenon of interest in the context of the study ($\mu_\delta = Cov(\delta,\xi) = Cov(\delta,x) = Cov(\delta,y) = Cov(\delta,\varepsilon_y) = Cov(\delta,\varepsilon_x) = 0$).

Further, $\rho_{zz'}$ indicates the percentage of variance of z that is attributable to the true value ξ. It can be simply thought of as the composite reliability (e. g., Cronbach α; Homburg and Giering 1997) of the control variable. If the control is measured using only one item, then $\rho_{zz'}$ is similar to what is typically referred to as indicator reliability (e. g., Klarmann and Homburg 2018).

To understand how the less-than-perfect reliability of the control affects the estimation of the effect of x (!) on y, we continue to look at the case in which all variables involved have a standard deviation of 1. If ξ, δ, and z are to have a standard deviation of 1, this implies:

$$z = \sqrt{\rho_{zz'}}\xi + \sqrt{(1-\rho_{zz'})}\delta_z \quad (14)$$

As a result, the correlation between x and z becomes:

$$r_{xz} = Cov\left(\alpha_x + \gamma\xi + \varepsilon_x, \sqrt{\rho_{zz'}}\xi + \sqrt{1-\rho_{zz'}}\delta_z\right) = \gamma\sqrt{\rho_{zz'}} \quad (15)$$

This result is, of course, the classical measurement attenuation result (e. g., Bollen 1989, p. 157): Random measurement error leads to correlations between observed variables that are smaller than the true correlations. Of importance is that measuring the control variable with error will also affect the estimate for the effect of x on y – even if both variables are not measured with error. In particular, by entering (14) and (15) into (12), we get:

$$b_1 = \frac{\beta_1 + \beta_2\gamma + Cov(x,\varepsilon_y) - \left(\beta_2\sqrt{\rho_{zz'}} + \beta_1\sqrt{\rho_{zz'}}\gamma + \sqrt{\rho_{zz'}}Cov(z,\varepsilon_y)\right)\sqrt{\rho_{zz'}}\gamma}{1-\rho_{zz'}\gamma^2}$$

$$= \beta_1 + \frac{1-\rho_{zz'}}{1-\rho_{zz'}\gamma^2}\beta_2\gamma + \frac{Cov(x,\varepsilon_y)}{1-\rho_{zz'}\gamma^2} - \frac{\rho_{zz'}\gamma Cov(z,\varepsilon_y)}{1-\rho_{zz'}\gamma^2} \tag{16}$$

This result shows that b_1 will be a biased estimate of the causal effect of x on y if the control z measures ξ with error. We cannot say whether the bias is positive or negative, as the direction depends on the causal pattern that is confounded (i. e., $\beta_2\gamma$). However, the bias is largest when $\rho = 0$, whereas the bias disappears when $\rho = 1$. Hence, only control variables that are measured well will improve the causal interpretability of results. It is worth noting that measurement error in z will also attenuate the estimated relationship between z and y.

2.1.4. Control variables and statistical testing

In most marketing research, the main interest does not lie in the parameter estimates as such, but in their statistical significance. Therefore, in this subsection we analyze how the common t-test for regression coefficients is affected if a common cause of x and y is omitted from the regression model. We do so without considering measurement error.

The starting point for hypothesis testing is the precision of the parameter estimates, measured through the standard error. In the bivariate case (and under the assumption of normal errors and homoskedastaticity), the standard error is defined as (e. g., Cameron 2006):

$$s.e._b = \sqrt{\frac{Var(e)}{(n-1)Var(x)}}. \tag{17}$$

As Var(x) will not be affected by endogeneity, the question is whether endogeneity affects the residual variance Var(ε). The residual variance is that part of the variance of y that is not explained by the regression equation. This is Var(ε) in the true bivariate model linking x to y:

$$Var(y) = Var(\alpha + \beta_1 x + \varepsilon)$$
$$Var(y) = \beta_1^2 Var(x) + Var(\varepsilon) + 2\beta_1 Cov(x,\varepsilon) \tag{18}$$
$$Var(\varepsilon) = Var(y) - \beta_1^2 Var(x) - 2\beta_1\beta_2\gamma Var(z)$$

With regard to OLS estimates, we know b from equation (9) when there is an omitted common cause. With that knowledge we can derive how Var(e) and Var(ε) are related:

$$Var(y) = Var(a + bx + e)$$
$$Var(y) = b^2 Var(x) + Var(e)$$
$$Var(e_y) = Var(y) - \left(\beta_1 + \beta_2\gamma\frac{Var(z)}{Var(x)}\right)^2 Var(x)$$

$$Var(e) = Var(y) - \beta_1^2 Var(x) - 2\beta_1\beta_2\gamma Var(z) - \beta_2^2\gamma^2\frac{Var(z)^2}{Var(x)} \tag{19}$$
$$Var(e) = Var(\varepsilon) - \beta_2^2\gamma^2\frac{Var(z)^2}{Var(x)}$$

From (19), we can see that the estimated residual variance in the model without the control variable will always be smaller than what it should be (Ebbes et al. 2017, p. 5). This result is to be expected. Since OLS minimizes residual variance by design, any solution that differs from the OLS solution must have a larger residual variance. Consequently, standard errors with omitted variables will be too small.

However, how the smaller standard errors affect hypothesis testing in total is unclear. In particular, the common t-test statistic to test hypotheses of the type H_0: $b = 0$ with H_1: $b \neq 0$ is defined as $t = b/(s.e._b)$. While the denominator will be too small, whether the numerator is smaller or larger than it should be will depend on the direction of the bias. As a result, we can only learn that null-hypothesis-significance testing can no longer be trusted for causal effects if common causes are omitted from the model.

2.1.5. Dealing with unobserved common causes

In everyday marketing research, situations often arise where the common cause is unobserved – that is, researchers do not have access to any measure of it. Importantly, common causes cannot be unobservable per se. In fact, any discussion of phenomena that cannot be measured at least in principle is meaningless (e. g., Carnap 1931).

Hence, in this section we consider cases where the context prevents researchers from having data on the required control variable. These situations mainly arise when researchers are using secondary data for their research, such as archival data or data collected through web crawlers. These situations are less likely to occur when researchers collect data themselves, particularly in survey research (or in experiments where randomization is expected to be incomplete). However, even in these contexts, reviewers may ask that the model include an additional causal control that was not included in the questionnaire. In this case, the methods presented here are helpful.

Quite a number of approaches to deal with unobserved common causes have been identified in the literature. We provide an overview of four approaches that we perceive to be especially relevant:

- Instrumental variables
- Fixed-effects regression
- Partitioning the distribution of error variance
- Sensitivity analysis.

The use of *instrumental variables* may be the most established approach when no data is available on the required control variable. Briefly, instead of regressing x on y, another variable – called the "instrument" – is regressed on y (for more extensive explanations, see Ebbes et al. 2017 and Morgan and Winship 2015). If this instrument is strongly correlated with x (relevance condition), but at the same time uncorrelated with z (exogeneity condition), the causal effect of x on y can be estimated accurately.

An important limitation of instrumental variable techniques is that whether the instrument is correlated with the omitted variable cannot be tested. Instead, it needs to be argued that this is the case, which may create some ambiguity. Moreover, many potentially exogenous instruments in marketing are not relevant (Rossi 2014), which leads to effect estimates and standard errors – neither of which can be trusted. In fact, if instruments are „only weakly correlated" with x and they are "even slightly endogenous," then estimates based on these instruments "are more biased and more likely to provide the wrong statistical inference than simple OLS estimates that make no correction for endogeneity" (Larcker and Rusticus 2010, p. 187). As a consequence, good instruments are not easy to identify.

Approaches that partition the distribution of the error variance of x are unique in the sense that they do not require variables beyond x and y to analyze and control for endogeneity (for overviews, see Ebbes et al. 2009 and Park and Gupta 2012). The general (and simplified) idea behind these approaches is that assumptions are possible about the distribution of the regression residuals in the true model (e. g., that they are normally distributed and/or homoskedastic). If a distribution is observed that differs from the assumed distribution, the difference is attributed to endogeneity. In turn, estimating the degree of endogeneity and accounting for it in estimating the relationship between x and y become possible – using techniques such as copulas (Park and Gupta 2012), higher moments (Erickson and Whited 2002), or latent instrumental variables (Ebbes et al. 2005).

For marketing researchers, these approaches are quite helpful in dealing with unobserved common causes. However, they do require a certain maturity of the field in that they replace the exogeneity assumption with an untestable assumption about the distribution of the true model residuals.

Fixed effects regression can be used with panel data – that is, data that include observations at multiple points in time for each case in the sample. This methodological staple from econometrics (e. g., Cameron and Trivedi 2005) allows researchers to separate different components of the error term. In particular, in fixed effects regression a case-specific error component is estimated. This error component will reflect the effect of all common causes to x and y that are left out of the model and (importantly) that are constant over time. Variables that are not constant over time still need to be included as control variables.

The purpose of *sensitivity analysis* is to be able to make statements about what types of omitted variables would alter the results of an analysis. That is, instead of trying to correct an estimated coefficient for an unobserved variable that is potentially a common cause of x and y, "one can determine what patterns of correlations result in estimates of this coefficient that are substantially different from the estimate originally obtained" (Mauro 1990, p. 317).

Mauro (1990) proposes an analytical approach, by which researchers determine at which point the correlations between x and z and between y and z are too substantial for the results to hold. With this approach, researchers need to provide extensive contingency tables that show the sensitivity analysis. Frank (2000) and Pan and Frank (2003) seek to lessen this complexity by reducing the sensitivity analysis to a single measure, namely the product of r_{xz} and r_{yz}. Their idea is to determine at which magnitude of $r_{xz} \times r_{yz}$ results change. Smaller values indicate less robustness to an unobserved omitted cause. VanderWeele (2015, pp. 66–97) extensively discusses methods that operationalize sensitivity as differences in the outcome variable.

In sum, quite a few promising approaches exist for dealing with situations in which researchers cannot directly observe common causes of x and y. Importantly, these approaches cannot outperform control variables that are validly measured. If data on a potential common cause are available, their inclusion in the model is preferable to all of these methods. Otherwise, these methods can be quite helpful. Given their limitations, they are particularly effective when combined. For instance, Papies et al. (2016) suggest that instrumental variable-free approaches using the partition of the distribution of error variance can well complement the instrumental variable approach to check the robustness of the findings.

2.1.6. The backdoor criterion

While the previous sections followed the dominant methodological paradigm of econometric analysis, requirements for control variables have also been formally derived under a different paradigm: causal graph analysis (Pearl 2009). Although related to structural equation modeling (Bollen and Pearl 2013), causal graph analysis completely separates estimation from identification (Elwert and Winship 2014). That is, the starting point is not an estimation method (like OLS in section 2.1.1). Instead, a more general understanding of causal identification is used: A causal effect is identified if this effect can be detected at all, given a sufficiently large dataset (Morgan and Winship 2015, p. 78). Causal graph analysis also does not require any idea about the functional form of the investigated relationships.

Causal graph analysis proceeds by analyzing the properties for "directed acyclic graphs" (Morgan and Winship 2015 provide an accessible introduction). Such graphs (*Fig. 1* shows examples) consist of nodes (the variables,

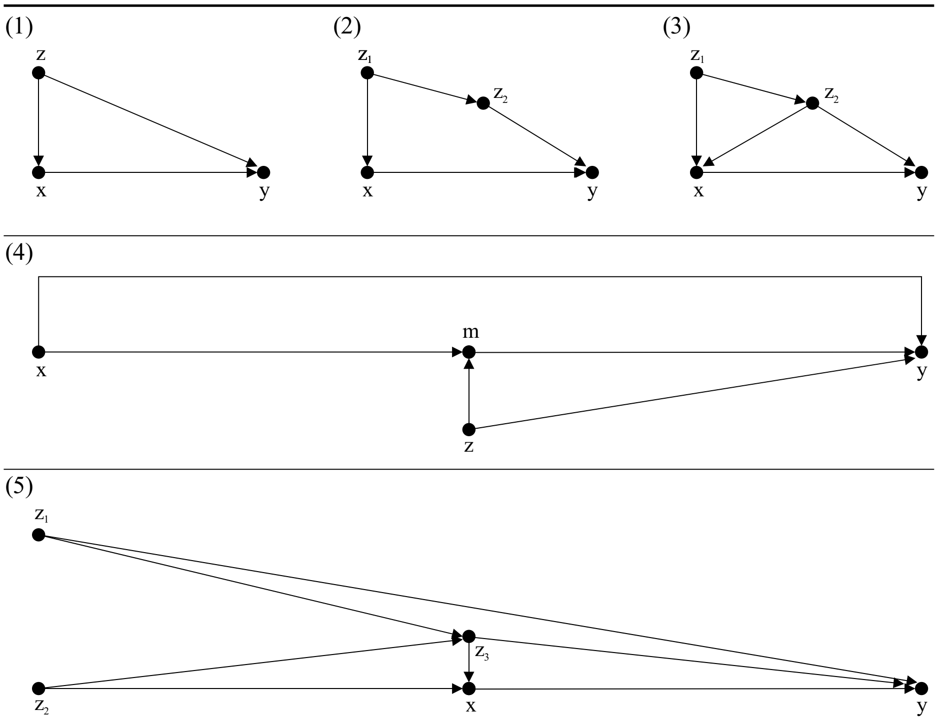

Fig. 1: Illustrations of Identification Analysis Using Causal Graphs

visualized as dots) and edges (the relationships, visualized as arrows). In causal graph analysis only directed graphs are considered. That is, one variable is always „the parent" of an effect (i. e., the origin), and another variable is the "descendant." Using arrows to represent the edges, the direction of these effects is made clear, with the arrow pointing from the parent to the descendant. The graphs considered are also acyclic, as one variable cannot be its own parent, either directly or indirectly. Additionally – given this tool's background in artificial intelligence research – a number of technical assumptions underlie these graphs (Pearl 2009), but can be ignored for the purposes of this paper.

A key result of this far-reaching research paradigm (according to Google Scholar, Pearl 2009 has been cited over 10,000 times so far) is the "backdoor criterion" to determine whether a causal effect is identified. Importantly, this criterion requires that researchers map out the entire nomological network of their research field in a causal graph like those shown in *Fig. 1*. This graph then allows identification of the required control variables. Specifically, the backdoor criterion is defined as follows: „Given an ordered pair of variables (X,Y) in a directed acyclic graph G, a set of variables Z satisfies the backdoor criterion relative to (X,Y) if no node in Z is a descendant of X, and Z blocks every path between X and Y that contains an arrow into X" (Pearl et al. 2016, p. 61).

Note that this criterion considers all possible paths from x to y, even if some edges in the graph in this path contain arrows pointing in the other direction. As in our econometric analysis in section 2.1.2, it is easy to verify that including z in the model in the first panel in *Fig. 1* satisfies the backdoor criterion. If z is included, every backdoor path between x and y is blocked, implying that by controlling for z, the causal effect of x on y is identified. In the second panel, the causal effect is identified if we control for z_1, z_2, or both. The third panel describes a situation discussed earlier: z_2 blocks all paths between x and y, even if it has a common cause with x that is not included in the model.

The model in the fourth panel is more complicated but is important, because it shows the classical mediation model (e. g., Zhao et al. 2010). The mediator m is a descendant of x. Hence, according to the backdoor criterion, controlling for m will no longer causally identify the direct effect of x on y. The reason is that m might share common causes with y that may in turn also influence the relationship between x and y in this model. This also motivates the advice discussed above not to control for post-treatment variables. Elwert and Winship (2014, pp. 44–45) provide an extensive discussion of this issue. To close all backdoor paths in the fourth panel, z also needs to be included as a control. Importantly, this requirement extends to experimental research, where only x is randomized, but not m (Pieters 2017 synthesizes the literature).

The model in the fifth panel of *Fig. 1* highlights the problem created by including control variables that themselves share common causes with y (see Morgan and Winship 2015, p. 110 for a more extensive discussion). In the example, this control variable is z_3. The causal effect of x on y is not identified if only z_3 is a control. According to the backdoor criterion, in this situation three sets of control variables identify the causal effect (Pearl et al. 2016, p. 61): $\{z_1,z_3\}$,$\{z_1,z_2,z_3\}$, and $\{z_2,z_3\}$.

Using the backdoor criterion to identify a set of control variables in applied marketing research projects will, of course, not be as straightforward as the examples in *Fig. 1*, as the relevant nomological networks are likely to be much more complicated. In this context identifying the relevant control variables also becomes much more complex. However, software approaches such as DAGitty are available to analyze graphs (Textor, Hard, and Knüppel 2011).

While applied researchers in particular will regularly lack knowledge of the causal graph required for this task, causal graph analysis can still be very helpful in applied marketing research. First, it clearly separates identification analysis (conceptual work) from model estimation (statistical work). As we state numerous times throughout this paper, a best practice in control variable use is to address control variable selection in the conceptual part of a manuscript. Second, causal graph analysis clarifies situations in which control variables increase bias instead of reducing it (e. g., Elwert and Winship 2014, p. 36). In particular, controlling for outcomes of x or variables that are outcomes of x and y („colliders") will make causal inference more difficult, not easier.

2.2. Control variables for improving the precision of estimated effects

In addition to including control variables in a model to improve the causal interpretability, adding variables to a model that are not evoked by the hypotheses can be done for a second plausible reason: to increase the precision of estimated effects – that is, to reduce the variability of estimates by decreasing their standard errors. This goal can be achieved by adding variables to the model that are causes of y, but not of x.

As we have seen before, adding these variables to the model will not affect the estimate of the effects, at least in larger samples. In small samples, despite x and y being uncorrelated in theory, sampling error might make them slightly correlated, which implies some effect on β_1. Imbens and Rubin (2015, pp. 122–125) provide a detailed discussion of this matter. The reduction of the standard error can be easily verified by looking at the formula for the standard error in a regression model with two independent variables, x and z. In such a model, the standard error of β_1 is (Cohen et al. 2003, p. 86):

$$s.e._{\beta_1} = \frac{\sqrt{Var(y)}}{\sqrt{Var(x)}} \sqrt{\frac{1}{1-r_{xz}^2}} \sqrt{\frac{1-R_y^2}{n-3}} \tag{20}$$

Importantly, (20) makes evident that the standard error is linked to the R^2 of y – that is, the amount of variance in y explained by x and z together. The higher the R^2, the smaller the standard error. Hence, by explaining more variance in y, the estimate of β_1 becomes less variable across samples. Intuitively, what is often referred to as „statistical noise" is reduced in y, which makes easier the estimation of the focal effect of x on y.

Equation (20) also makes evident that this increase in precision will be at least partly lost if x and z are correlated. Then, the causal relationship and the causal status of z become relevant again, to ensure that the results are valid. Only if no correlation exists between x and z can adding z to the model be motivated through increased precision.

Since adding controls to improve precision requires uncorrelatedness, this approach is not surprisingly used most often in research using randomized experiments. In these experiments, x and z are uncorrelated by design (at least they should be). A good example for this approach is experimental research on consumers' self-control and indulgent consumption. Here, large shares in the variances of the outcomes of interest (e. g., the likelihood of choosing an unhealthy over a healthy option; social consequences of indulgence or restraint) are quite often explained simply by the participants' gender and the extent to which they are hungry at the time of the experimental study, which is why gender and hunger are frequently included as control variables (e. g., Haws and Winterich 2013; Lowe and Haws 2014).

3. Practical perspective and recommendations

In this section, we review the descriptive empirical literature on the use of control variables. By combining insights on applied behavior and the theoretical knowledge from section 2, we derive recommendations for marketing researchers. *Tab. 1*, which is based on an analysis of published articles in leading journals (mostly in the domain of organizational behavior), lists the studies that describe and discuss the correct use of control variables

Even though these studies differ in focus and provide recommendations at different specificity, the authors' opinions on what should be done generally converge. Synthesis of the observations made in the review articles mentioned above reveals much room for improvement regarding control variable use in research practice.

In what follows, we extract from the literature key questions marketing researchers might ask themselves in what we call the "control variable process," and we suggest when these questions should be addressed within the general research process. Notably, in this framework, control variable considerations start in the conceptualization phase of research – but only after the research question is finalized. This timing reflects what distinguishes control variables from the focal variables in a framework: They are not evoked by the research question itself.

The control variable process we propose has six steps and is depicted in *Fig. 2*. The process starts with control variable selection and ends with the interpretation of research results. In the following, we go through these six steps and discuss important questions that arise in this phase. Wherever possible, we support our arguments by providing descriptive statistics about actual control vari-

Authors (Year)	Journal	Base	Main contribution
Becker (2005)	*Organizational Research Methods*	60 empirical articles published during 2000-2002 in the *Academy of Management Journal, Journal of Applied Psychology, Journal of Management,* and *Strategic Management Journal*	12 recommendations regarding the treatment of control variables
Breaugh (2008)	*Human Resource Management Review*	59 empirical articles published during 2005-2006 in the *Journal of Applied Psychology* and *Personnel Psychology*	Examples from the literature and discussion of the confusion that including control variables can cause
Carlson and Wu (2012)	*Organizational Research Methods*	162 empirical articles published during 2007 in the *Academy of Management Journal, Journal of Applied Psychology*, and *Strategic Management Journal*	7 recommendations regarding the treatment of control variables
Atinc et al. (2012)	*Organizational Research Methods*	812 empirical articles published during 2005-2009 in the *Academy of Management Journal, Journal of Applied Psychology, Journal of Management,* and *Strategic Management Journal*	Several specific recommendations for both authors and reviewers/editors
Bernerth and Aguinis (2016)	*Personnel Psychology*	580 empirical articles published during 2003-2012 in the *Academy of Management Journal, Administrative Science Quarterly, Journal of Applied Psychology, Journal of Management,* and *Personnel Psychology*	Decision-making tree summarizing sequential steps in the process of selecting control variables
Bernerth et al. (2018)	*Journal of Management*	290 empirical articles published during 2003-2014 in the *Academy of Management Journal, Administrative Science Quarterly, Group & Organization Management, Journal of Applied Psychology, Journal of Business and Psychology, Journal of Organizational Behavior, Journal of Management, The Leadership Quarterly, Organizational Behavior and Human Decision Processes*, and *Personnel Psychology*	Detailed recommendations for justifying selected control variables

Tab. 1: Selection of Empirical Literature on Control Variable Use

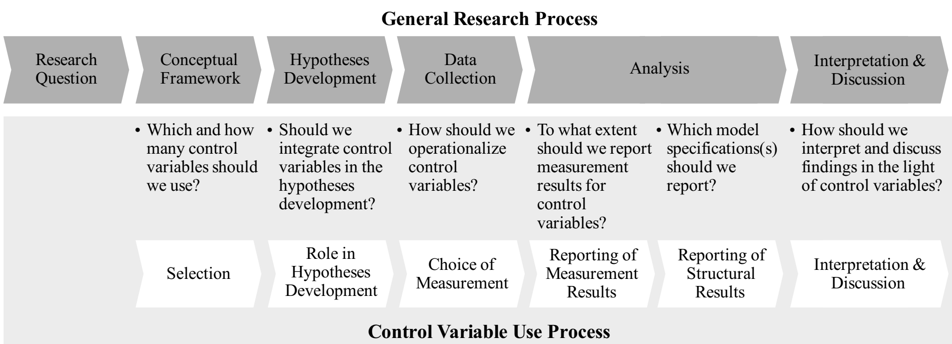

Fig. 2: Important Questions Arising in the Control Variable Use Process

able use in research practice. In each section, we develop specific recommendations, with a focus on feasibility (e. g., regarding questionnaire length and the review process). The recommendations are summarized in *Tab. 2* at the end of this paper.

3.1. Control variable selection

Atinc et al. (2012) find that in a sample of more than 800 management studies, almost one-fifth (18.2 %) of all studies under investigation provide no rationale for the inclusion of control variables. Only about half of the studies (48 %) provide a sound theoretical reason or prior empirical evidence. We have the impression that a sample of empirical studies in marketing would lead to similar results. However, consistent with previous recommendations (and in line with section 2 on control variable theory), researchers should make a compelling case – ideally based on logic and theory – for which control variables need to be included (Spector and Brannick 2011).

In particular, we strongly recommend separating causal identification from model estimation (in line with the literature presented in section 2.1.6). This recommendation implies that in any research project aiming at causal inference, control variable selection needs to be a key part of the conceptualization phase. Consistent with Becker et al. (2016), we recommend that control variables become part of the conceptual model guiding the research.

We suggest that researchers sketch a comprehensive causal graph around the focal variables. The result is a graphic depiction of all potentially relevant variables and their assumed interrelations based on considerations of theory, prior empirical evidence, logic, and plausibility. This causal graph allows the identification of all common causes, which if omitted would create bias in the effect of interest. The graph also makes possible the use of the backdoor criterion for control variable selection. The software DAGitty (Textor, Hard, and Knüppel 2011) allows for an automated detection of correct sets of control variables given a causal graph.

The applied problem, of course, is that in many research projects – especially those seeking to uncover new phenomena – important aspects of this causal graph may be based not on sound knowledge but on guesswork. The reviewers of control variable practice observe a tendency to include too many rather than too few control variables.

They attribute this tendency to researchers' misconceptions that the inclusion of control variables "purifies results," is "playing it safe," and leads to a more "conservative" hypothesis testing (Bernerth and Aguinis 2016; Carlson and Wu 2012; Meehl 1971; Spector and Brannick 2011). This interpretation has led to the rule of thumb specifying "when in doubt, leave them out," stemming from the fear that adding more and more control variables may render the interpretation of results more difficult (Becker et al. 2016; Carlson and Wu 2012). Our view is that this advice might go a little too far. As discussed in section 2.1.1, excluding important controls from the model creates more bias than including irrelevant ones – at least if these irrelevant variables are uncorrelated with x. This view is in line with Bernerth et al. (2018), who conclude that when it comes to the inclusion of control variables, not less is more, but "less of the nontheoretical type is more" (p. 24). Of course, this conclusion implies that researchers need to be up to the task of adequately interpreting and discussing the results (see section 3.6).

We recommend basing control variable selection for improving causal inference on the following selection criteria:

- *Completeness*: All relevant common causes between x and y should be controlled for.
- *Cause of x*: All control variables should plausibly affect x.
- *Cause of y*: All control variables should plausibly affect y.
- *Exogeneity with regard to y*: The relationship between z and y should not potentially be explained by common causes of z and y. Otherwise, these variables also need to be included as controls.
- *No outcome of x*: To causally identify a relationship between x and y, control variables should not be caused by x.

Every control variable should be introduced in a way that explains how these five criteria are met. This requirement is consistent with advice in the literature to elaborate on a theoretical rationale in a comprehensible manner and allow the rationale the manuscript space it deserves (e. g., Becker 2005). Authors are also advised to support their rationale by citing empirical evidence from prior research (ideally from meta-analyses) and by including a statement regarding the expected direction of effect of the control variable (e. g., Becker 2005).

Against this backdrop, we suggest that authors explicitly include the control variables in their theoretical model and provide at least a short comprehensible justification based on theory and empirical findings in the theoretical or conceptual background section. This motivation should also provide a clear definition of the control variable. Homburg, Jensen, and Hahn's (2012a) article serves as a good example for this approach.

3.2. Control variables in hypotheses development

Control variables are rarely mentioned in the hypotheses section, let alone in the formal hypotheses. This absence is sometimes seen as problematic because the analyses typically include control variables, creating the impression that a mismatch exists between the relationships that are hypothesized and those that are actually tested. Consequently, some commentators (e. g., Becker et al. 2016; Spector and Bannick 2011) call for inclusion of control variables in the hypotheses if they are included in the analysis (e. g., "H_1: Controlling for corporate branding, CMO presence has a positive effect on firm performance").

However, as prior work suggests (Atinc et al. 2012; Bernerth et al. 2018), this advice is presently only very rarely followed by authors (3.6 % and < 1 %, respectively). The reason is understandable. Most importantly, even though control variables are added to the model to improve causal inference, the relationship of interest is most likely bivariate (the causal effect of x on y). That is, isolating a causal relationship makes a multivariate analysis necessary, but the hypothesis is about the variables x and y. Additionally, unconditionally following the advice to acknowledge control variables will quickly lead to very long and unreadable hypotheses if the set of controls is larger than, say, three.

Therefore, we suggest that authors refrain from including control variables in their hypotheses. Nevertheless, the causal network surrounding the relationship of interest should be evoked in the argument leading to the hypothesis. Moreover, we suggest including a brief acknowledgement at the beginning of the hypothesis development section, stating that the bivariate character of the hypotheses reflects the expected causal relationship are only identified when holding the (previously introduced and justified) control variables constant.

3.3. Control variable measurement

Researchers usually invest considerable time and effort in ensuring that the focal variables are not only properly defined and conceptualized but also reliably and validly measured (Homburg and Giering 1996). Researchers typically take less care about measurement issues when it comes to control variables. As section 2.1.3 described, a control variable that is measured unreliably will not completely remove the bias from the relevant coefficient (see also Bernerth and Aguinis 2016). Hence, researchers should adopt the same high standard of measuring control variables as they adopt for their focal variables, especially using pretesting, multi-item scales, and – if possible – multiple informants (Homburg et al. 2012c).

Researchers are also advised not to use proxy control variables (i. e., surrogates for other, meaningful control variables, such as using firm size as a proxy for the degree of formalization), as the relationship strength between the proxy control variable and the dependent variable often differs from the relationship strength between the meaningful control variable and the dependent variable to an unknown extent (Becker et al. 2016; Breaugh 2008). In fact, a proxy control variable may relate to other focal variables in a way that the more meaningful control variable does not, and hence, "control for a host of unintended variables that have substantive effects that the researcher does not wish to remove" (Becker et al. 2016, p. 161).

Last, while some authors may deem a reliable measurement of control variables as unnecessary in the first place, in other cases the reporting of which scales were used to measure control variables may have fallen prey to the researchers' efforts to stick to the journal's page limits. In 10 % of all studies analyzed by Atinc et al. (2012), no measurement information whatsoever was provided for control variables.

We strongly agree with previous commentators that measurement of control variables needs the same level of attention as the measurement of focal variables and that proxies should be avoided. Fortunately, measuring control variables with multi-item scales and reporting the appropriate information about measurement adequacy should be easily defensible in the review process. Instead, feasibility may be threatened by the length of the questionnaire. Given the obvious adverse effects of an overly long questionnaire on response rates and answering behavior (especially for busy managers), including a multi-item measure for each control variable in a questionnaire is impractical. Diamantopoulos et al. (2012) and Fuchs and Diamantopoulos (2009) clarify in which cases single-item measures are sufficient, and we encourage authors to follow and cite their arguments. Whatever the choice is, we agree with prior commentators (e. g., Becker 2005) that authors should make fully transparent which scales were used.

3.4. Control variables and measurement reporting

Researchers' scant attention to control variables also often becomes apparent in the inadequate reporting of measurement results as far as they relate to control variables. Researchers frequently seem to adopt a two-tier approach to reporting the psychometric properties of control variables and their impact on variables beyond the focal variables. In 8 % of studies analyzed by Bernerth et al. (2018), reliability indices of the control variable measurements were not reported. Furthermore, in 10.5 % of all studies analyzed by Atinc et al. (2012), control variables were not included in the correlation table, and in 13.7 % of all studies, descriptive statistics of control variables were not reported.

The literature on best control variable practice is relatively clear on this issue: "Control variables deserve as much attention and respect as do independent and dependent variables" (Becker 2005, p. 286). Specifically, researchers should report to the same level of detail as they do for the focal variables – that is, report measurement information, include control variables in correlation tables, and present descriptive statistics (Becker et al. 2016).

We agree with literature proposing that authors should apply the same care for control variables as for focal variables when it comes to analysis and reporting. In line with Atinc et al. (2012), we encourage researchers to conduct confirmatory factor analysis for all control variables (if applicable), and report the resulting indices and psychometric properties including indicator reliability, composite reliability or Cronbach's α, and average variance extracted (Homburg and Giering 1996). Control variables should also be included in correlation/discriminant validity analyses (Fornell and Larcker 1981; Henseler et al. 2015).

3.5. Control variables and results reporting

The proscription of two-tier reporting of results also pertains to the results of the multivariate analysis and the actual hypotheses testing. Becker (2005) correctly points out that, from the reader's perspective, control variables could be focal variables in future studies, and without full knowledge about which control variables were included and what their effects are, replication studies are not effective. The question remains as to which is the „right" model specification to report in the results section.

Prior commentators suggest that authors run and report separate analyses and include results tables for different models in which control variables are included and excluded to demonstrate the control variables' impact on the relations between the independent variables and the dependent variable (Atinc et al. 2012; Becker 2005; Becker et al. 2016). However, in Bernerth et al's (2018) analysis, only 5 % of all studies followed this recommendation. Again, the reasons are understandable. If control variables are required for causal identification, results will differ between causally unidentified models (without any controls or with a wrong subset of controls) and identified models (with correct subsets of controls). While differences should be expected, readers and reviewers alike might mistake the differences for inconsistencies that limit the credibility of findings. Consider once more Panel 5 in *Fig. 1*. In this instance, the possible models are eight: one without controls, three with one control each, three with two controls each, and one including all controls. From the discussion in section 2.1.2, we know that only three of these eight models satisfy the backdoor criterion. If a researcher were to report all eight models, five would be misspecified, creating a strong imbalance with regard to the relationships revealed between x and y. The number of potentially wrong models will be much higher in larger causal graphs.

Consequently, authors should limit the reporting of alternative model specifications to a subset that is causally identified according to their own conceptual framework. This subset could for instance be a baseline model that includes only the control variables and a final model that also considers the relationships of interest. If whether certain control variables are required for causal identification is unclear, models with these variables included should be compared to models without the variables.

Recently, Simonsohn et al. (2015) have proposed an alternative way of comparing model results: specification curve analysis, which analyzes all possible model specifications (i. e., with all possible combinations of control variables included in or excluded from the model). Then, hypothesis tests are done across results from all models. Nevertheless, the authors warn researchers to include only model specifications from a subset they consider theoretically valid: "Researchers do not need to estimate specifications they consider redundant, and certainly not specifications they consider invalid" (p. 3). Again, this caution points to the necessity of having a sound conceptual background guiding control variable inclusion.

3.6. Control variables in the interpretation and discussion of results

Authors may report the effects of the control variables on the dependent variable, but often ignore these effects in the interpretation and discussion of the results. As Atinc et al.'s (2012) work suggests, this disregard is quite typical, as 73 % of all studies do not mention control variables in the discussion section. Consistent with other commentators, we recommend improvement of this situation. Most importantly, control variables are added to a model to improve the causal interpretability of results. The discussion section must consider whether causal inferences are indeed possible given the data, variables, and results. Also advisable is some elaboration on the extent to which conceptual considerations and the underlying causal graph are supported by the data. Here are some questions that could be addressed: Are the control variables related to the independent variables and the dependent variables? What does this relationship imply for future research? Does any evidence show that control variables are missing? Were some controls ineffective, owing to measurement error?

We also recommend openly discussing the extent to which control variables drive the findings. Such a discussion is not disreputable but expected, important, and potentially interesting. The more authors follow our advice to make the control variables an essential part of the manuscript's story and highlight the multivariate character of the analysis early on, the more easily they will be able to follow this recommendation. Ideally, authors should make an effort to provide specific situations to which the results of the study do and do not generalize. Again, as some reviewers may perceive inability to generalize as undermining the paper's contribution, more risk-averse authors should be cautious about generalizing results and avoid misleading claims regarding generalizability.

Notably, the literature on control variable use from the management domain raises two concerns about the interpretation of control variable use. Researchers could consider these concerns in their discussion of results based on the use of control variables. First, prior research on control variable use observes that researchers generally ignore the multivariate character of their analysis in discussing their findings. They suggest emphasizing in the discussion that the interpretation of the estimated coefficient refers to situations when the controls are held constant (Becker et al. 2016, Breaugh 2008).

Second, drawing on a handbook chapter by Meehl (1970), Becker et al. (2016) fear that the isolated effect measured through regression coefficients in models with many variables may not generalize well, because in practice phenomena are strongly interrelated. We at least partly disagree. Managers are likely to be interested in isolating the effect of individual measures. Not surprisingly, in his seminal book on causality, Pearl (2009) defines and uses what he calls a "do operator." Control variables can give findings a causal interpretation that produces much more actionable implications than simple correlations. By controlling for common causes, researchers hold these variables constant. If a researcher finds that a marketing variable has an impact on performance while holding firm size, marketing expenditures, firm age, and industry constant, managers infer that the marketing variable "works" regardless of firm size, marketing expenditures, firm age, and industry. In our eyes, this interpretation strengthens the implications rather than weakening them.

Finally, control variables are, of course, not a general remedy to all problems related to causal inference. Researchers need to acknowledge and discuss other endogeneity issues.

4. Concluding directions for researchers

The overarching goal of our study was to develop recommendations for good control variable use in marketing research. We were able to do so by combining insights from econometrics, test theory, causal graph analysis, and descriptive analyses of control variable practice. Our recommendations are summarized in *Tab. 2*, with reference to the six steps of the control variable process described in section 3. We believe that our recommendations provide feasible, and hence valuable, guidelines for better control variable use by marketing researchers. We hope that future research (and reviewers) will embrace at least some of our recommendations to craft research that ultimately allows to derive actionable implications that managers (and editors) desire (e. g., Kumar 2016).

With the increasing emphasis on causal inference in marketing research, future research generations are likely to require even more knowledge about control variable use than what is currently available. We offer three more recommendations for future research that we believe could be particularly fruitful over and above the recommendations we provide in *Tab 2*.

Control variable use process stage	Recommendation	Rationale
Selection	1. Separate causal identification from model estimation and move it to the front end of the paper.	Ensures causal identification independent of the functional form of the relationships and assumptions required for model estimation.
	2. Start out by mapping the causal graph around the focal variables.	Reveals the relations between variables and indicates which should cause biased results if omitted.
	3. Make an effort to include control variables only if a sound rationale can be provided. If the evidence is unclear, include them and be sure to follow recommendation 12.	Follows the notion that too few control variables are potentially more harmful than too many.
	4. Include control variables in the conceptual framework (both text and figures)	Upgrades control variables to be of conceptual importance.
	5. Include at least a short motivating statement based on theory and empirical findings for each control variable. Justify completeness of control variables and show that effects on x and y are plausible.	Discourages random selection of control variables and provides reasonable explanation given the page constraints.
	6. Make sure there are no post-treatment controls and endogenous controls.	Avoids introduction of new biases for causal inference.
Role in Hypothesis Development	7. Clarify at the beginning of the hypotheses section that all hypotheses are stated assuming that control variables are held constant.	Stresses the multivariate character of the study sufficiently given the page limits.
Choice of Measurement	8. Use multi-item scales for latent control variables if appropriate.	Demonstrates sophisticated use of measurements.
	9. Avoid proxies.	Ensures that only conceptually meaningful control variables are included.
Reporting of Measurement Results	10. Conduct factor analyses for all control variables.	Ensures the accuracy of parameter estimates for independent variables.
Reporting of Structural Results	11. Report standard descriptive statistics for control variables and include control variables in correlation table	Enhances understanding of psychometric properties and replicability
	12. Estimate and compare valid models, especially if the inclusion of some control variables cannot be soundly argued conceptually.	Demonstrates the impact and importance of control variables for the results and the variance explained.
	13. Explicitly report the effects of control variables on the dependent variable in results tables.	Illuminates the impact of control variables on the dependent variable.
Interpretation and Discussion	14. Discuss strength of causal claims.	Ensures transparency of the role of control variables on the focal causal effect(s).
	15. Discuss what the results imply for the underlying causal graph.	Allows other researchers to make more educated choices with regard to control variables.
	16. Discuss whether other types of endogeneity could still be a problem.	Recognizes that control variables are no "silver bullet" to solving endogeneity problems.

Tab. 2: Recommendations

First, as has become evident throughout this paper, to causally identify a relationship between x and y based on observational data, knowledge of the causal graph surrounding these variables is essential. However, in many situations the researcher will lack access to empirical evidence in this regard. Even if meta-analyses in a given field are available, they tend to focus on bivariate relationships between variables. In fact, aggregating correlations is how meta-analyses primarily proceed (e. g., Eisend 2015). To improve control variable use in a given field, something resembling a "causal meta-analysis" is needed. Studies of this type would not focus on the bivariate correlations between two variables alone, but would seek to restore causal graphs surrounding the phenomena.

Second, uncertainty remains as to how to assess the robustness of results on the basis of control variables. On the one hand is the fear of false positive research, which increases the demands on researchers to show that their results are as robust as possible across multiple model specifications. On the other hand, is the possibility of false negative research, because an effect disappears in some potential control variable constellations that are misspecifications but are not recognized as such by reviewers. Specification curve analysis is a promising approach to better assess result consistency. It will be important, though, to develop some joint understanding about what degree of convergence of estimates across models is considered as support for a hypothesis – and what is not. This agreement will increase the transparency of control variable use and bring the field closer to causal conclusions, even when using observational data.

Last, we note that our development of recommendations for marketing researchers is based on critical reviews of control variable research in other domains. Even though we do not expect that current control variable practice in marketing deviates from control variable practice in organizational behavior, conceivably it may. We therefore acknowledge that our recommendations may neglect issues that are of specific concern in the marketing literature, and it would be interesting to qualitatively and quantitatively examine control variable practice in leading marketing journals, similar to the studies listed in *Tab. 1*.

References

Angrist, J. D. & Pischke, J.-S.(2015). *Mastering 'Metrics: The Path from Cause to Effect*, Princeton: Princeton University Press.

Atinc, G., Simmering, M. J., & Kroll, M. J. (2012). Control Variable Use and Reporting in Macro and Micro Management Research. *Organizational Research Methods*, 15(1), 57–74.

Becker, T. E. (2005). Potential Problems in the Statistical Control of Variables in Organizational Research: A Qualitative Analysis with Recommendations. *Organizational Research Methods*, 8(3), 274–289.

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior*, 37(2), 157–167.

Bernerth, J. B., & Aguinis, H. (2016). A Critical Review and Best-Practice Recommendations for Control Variable Usage. *Personnel Psychology*, 69(1), 229–283.

Bernerth, J. B., Cole, M. S., Taylor, E. C., & Walker, H. J. (2018). Control Variables in Leadership Research: A Qualitative and Quantitative Review. *Journal of Management*, 44(1), 131–160.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*, Hoboken: John Wiley & Sons, Inc.

Bollen, K. A. & Pearl, J. (2013). Eight Myths about Causality and Structural Equation Models. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research*, Dortrecht: Springer, 301–328.

Bono, J. E., & McNamara, G. (2011). Publishing in AMJ-Part 2: Research Design. *Academy of Management Journal*, 54(4), 657–660.

Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review*, 18(4), 282–293.

Cameron, A. C. (2006). Review of Bivariate Regression. Unpublished manuscript available at http://cameron.econ.ucdavis.edu/e240a/reviewbivariate.pdf.

Cameron, A. C. & Trivedi, P. K. (2005). *Microeconomics: Methods and Applications*, New York: Cambridge University Press.

Carlson, K. D., & Wu, J. (2012). The Illusion of Statistical Control: Control Variable Practice in Management Research. *Organizational Research Methods*, 15(3), 413–435.

Carnap, R. (1931). Überwindung der Metaphysik durch logische Analyse der Sprache. *Erkenntnis*, 2(1), 219–241.

Cohen, J., Cohen P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Ed., Mahwah: Lawrence Erlbaum.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3), 434–449.

Ebbes, P. Papies, D. & van Heerde, H. J. (2017). Dealing with Endogeneity: A Nontechnical Guide for Marketing Researchers. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of Market Research*, Cham: Springer, n/a.

Ebbes, P., Wedel, M. & Böckenholt U. (2009). Frugal IV Alternatives to Identify the Parameter for an Endogenous Regressor. *Journal of Applied Econometrics*, 24(4), 446–468.

Ebbes, P., Wedel, M., Böckenholt, Z., & 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(4), 365–392.

Edwards, J. R., & Bagozzi, R. P. (2000). Successful Change Programs Begin with Results. *On the Nature and Direction of Relationships between Constructs and Measures*, 5(2), 155–174.

Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden*, Weinheim: Beltz.

Eisend, M. (2015). Have We Progressed Marketing Knowledge? A Meta-Meta-Analysis of Effect Sizes in Marketing Research. *Journal of Marketing*, 79(3), 23–40.

Elwert F. & Winship, C. (2014). Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual Review of Sociology*, 40, 31–53.

Erickson, T. & Whited, T. M. (2002). Two-Step GMM Estimation of the Errors-In-Variables Model Using High-Order Moments. *Econometric Theory*, 18(3), 776–799.

Fornell, C. & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50.

Frank, K. A. (2000). Impact of a Confounding Variable on a Regression Coefficient. *Sociological Methods & Research*, 29(2), 147–194.

Fuchs, C., & Diamantopoulos, A. (2009). Using single-item measures for construct measurement in management research. *Business Administration Review*, 69, 195–210.

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge: Cambridge University Press.

Germann, F., Ebbes, P., & Grewal, R. (2015). The Chief Marketing Officer Matters! *Journal of Marketing*, 79(3), 1–22.

Habel, J., & Klarmann, M. (2014). Customer Reactions to Downsizing: When and how is Satisfaction Affected? *Journal of the Academy of Marketing Science*, 43(6), 768–789.

Haws, K. L., & Winterich, K. P. (2013). When Value Trumps Health in a Supersized World. *Journal of Marketing*, 77(3), 48–64.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135.

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). Prediction and explanation in social systems. *Science*, 355, 486–488.

Homburg, C., & Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte. *Marketing ZFP-Zeitschrift für Forschung und Praxis*, 18(1), 5–24.

Homburg, C., Jensen, H., & Hahn, A. (2012a). How to Organize Pricing? Vertical Delegation and Horizontal Dispersion of Pricing Authority. *Journal of Marketing*, 76(5), 49–69.

Homburg, C., Klarmann, M., & Staritz, S. (2012b). Customer Uncertainty Following Downsizing: The Effects of Extent of Downsizing and Open Communication. *Journal of Marketing*, 76(3), 112–129.

Homburg, C., Klarmann, M., Reimann, M., & Schilke, O. (2012c). What Drives Key Informant Accuracy? *Journal of Marketing Research*, 49(4), 594–608.

Imbens, G. W. & Rubin, D.B. (2015). *Causal Inference for Statistics, Social, and Bionedical Sciences: An Introduction*, New York: Cambridge University Press.

Kirk, R. E. (1995). Experimental design: Procedures for the behavioral sciences (3rd Ed.), Pacific Grove: Brooks/Cole.

Klarmann, M. & Homburg, C. (2018). Factor Analysis. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of Market Research*, Cham: Springer, n/a.

Koschate-Fischer, N., & Schandelmeier, S. (2014). A Guideline for Designing Experimental Studies in Marketing Research and a Critical Discussion of Selected Problem Areas. *Journal of Business Economics*, 84(6), 793–826.

Kumar, V. (2016). My Reflections on Publishing in Journal of Marketing. *Journal of Marketing*, 80(1), 1–6.

Larcker, D. F. & Rusticus, T. O. (2010), On the Use of Instrumental Variables in Accounting Research. *Journal of Accounting and Economics*, 49(3), 186–205.

Lowe, M. L., & Haws, K. L. (2014). (Im)moral Support: The Social Outcomes of Parallel Self-Control Decisions. *Journal of Consumer Research*, 41(2), 489–505.

Marketing Science Institute (2016), *Research Priorities 2016–2018*, Cambridge: Marketing Science Institute.

Mauro, R. (1990). Understanding LOVE (Left out Variables Error): A Method for Estimating the Effects of Omitted Variables. *Psychological Bulletin*, 108(2), 314–329.

McDonald, R. P. (1999). *Test theory: A unified approach.* Mahwah, Lawrence Erlbaum.

Meehl, P. E. (1970). Nuisance Variables and the Ex Post Facto Design. In M. Radner & S. Winokur (Eds.), *Analyses of Theories and Methods of Physics and Psychology*, Minneapolis: University of Minnesota Press, 373–402.

Meehl, P. E. (1971). High School Yearbooks: A Reply to Schwarz. *Journal of Abnormal Psychology*, 77(2), 143–148.

Morgan, Stephen L. & Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, 2nd Ed., New York: Cambridge University Press.

Pan, W., & Frank, K. A. (2003). A Probability Index of the Robustness of a Causal Inference. *Journal of Educational and Behavioral Statistics*, 28(4), 315–337.

Papies, D., Ebbes, P. & van Heerde, (2016). Addressing Endogeneity in Marketing Models. In P. S. H. Leeflang, J. E. Wieringa, T. H. A. Bijmolt, K. H. & Pauwels (Eds.), *International Series in Quantitative Marketing*, Cham: Springer, 581–627.

Park, S. & Gupta, S. (2012), Handling Endogenous Regressors by Joint Estimation Using Copulas. *Marketing Science*, 31(4), 567–86.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd Ed., New York: Cambridge University Press.

Pearl, J. (2014), Comment: Understanding Simpson's Paradox. *The American Statistician*, 68(1), 8–13.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer.* Chichester: John Wiley & Sons.

Pieters, R. (2017). Meaningful Mediation Analysis: Plausible Causal Inference and Informative Communication. *Journal of Consumer Research*, 44(3), 692–716.

Rossi, P. (2014). Even the Rich Can Make Themselves Poor: A critical Examination of the Use of IV Methods in Marketing. *Marketing Science*, 33(5), 655–672.

Pratt, J. W. & Schlaifer, R. (1988). On the Interpretation and Observation of Laws. *Journal of Econometrics*, 39(1–2), 23–52.

Sande, J. B. & Ghosh, M. (2018), Endogeneity in survey research. *International Journal of Research in Marketing*, accepted for publication at January 22nd, 2018.

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3), 289–310.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology. *Psychological Science*, 22(11), 1359–1366.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A Key to The File Drawer. *Journal of Experimental Psychology: General*, 143(2), 534–547.

Simonsohn, U., Simmons, J. P. and Nelson, L. D. (2015). Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications (November 24). Available at SSRN: https://ssrn.com/abstract=2694998 or http://dx.doi.org/10.2139/ssrn.2694998.

Spector, P. E., & Brannick, M. T. (2010). Methodological Urban Legends: The Misuse of Statistical Control Variables. *Organizational Research Methods*, 14(2), 287–305.

Steyer, R. & Schmitt, T. (1994). The Theory of Confounding and Its Application in Causal Modeling With Latent Variables. In A. von Eye, A. & C. C. Clogg (Eds.), *Latent Variables Analysis: Applications for Developmental Research*, Thousand Oaks: Sage, 36–67.

Textor, J., Hardt, J., & Knüppel, S. (2011), DAGitty: A Graphical Tool for Aanalyzing Causal Diagrams, Epidemiology, 22(5), 745.

Thoemmes, F., Rose, N. (2014). A cautious note on the use of auxiliary variables for missing data problems. *Multivariate Behavioral Research*, 49:5, 443–459.

VanderWeele, T. J. (2015). *Explanation in Casual Inference: Methods for Mediation and Interaction*, Oxford: Oxford University Press.

Zhao, Xinshu, John G. Lynch Jr., & Qimei Chen (2010), Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, 37(2), 197–206.

Keywords

Control Variables, Covariates, Cross-Sectional Data, Endogeneity, Observational Data, Survey Research.