# Control Variables in Interactive Models

**Ed deHaan**
University of Washington
edehaan@uw.edu

**Robbie Moon**
Georgia Institute of Technology
robbie.moon@scheller.gatech.edu

**Jonathan Shipman**
University of Arkansas
jshipman@walton.uark.edu

**Quinn Swanquist**
University of Alabama
qtswanquist@cba.ua.edu

**Robert Whited**
North Carolina State University
rwhited@ncsu.edu

February 2023

**Abstract:** Accounting studies often examine whether the relation between *X* and *Y* varies with a moderating variable, *M*, by including an interactive term, *X*×*M*, in a regression. We provide plain-English guidance on why, how, and when to use control variables, *Z*, in interaction tests. A simulation and simple descriptions demonstrate how interacted controls affect coefficient estimates and interpretations. In particular, we demonstrate how including *Z* without an accompanying interaction of *X*×*Z* or *M*×*Z* generally does not eliminate the confounding effect of *Z* on *X*×*M*. We conclude with guidance for future research.

-------

* Stata code to produce the simulations in this paper is available at: github.com/ed-dehaan/Interacted_Controls. Thanks to Sarah McVay, Dan Taylor, and John Wertz for helpful discussions. This is a working draft. Comments are welcome.

## 1. Introduction

Accounting studies often examine whether the relation between two variables, treatment ($X$) and outcome ($Y$), varies with a moderating variable ($M$). For example, we might start by examining the effect of earnings surprises ($X$) on stock returns ($Y$) by estimating the average "earnings response" across all observations in a sample. Because it is unlikely that the relation between earnings and returns is the same for all firms, we could perform a cross-sectional test of whether the strength of earnings responses varies with a characteristic such as firm size ($M$). Researchers frequently execute such tests using an ordinary least squares regression model with an interaction term ($X \times M$):

$$Y = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 X \times M + \varepsilon \quad (1)$$

Here, $M$ is a "moderating variable," meaning that it explains the conditions under which the relation between $X$ and $Y$ differs in strength or sign (Jollineau and Bowen 2022). $\beta_3$ reflects the incremental marginal effect of $X$ on $Y$ for a one-unit increase in $M$ and is the coefficient of interest in a "cross-sectional" or "interactive" test.

In observational settings, studies often recognize that another variable, $Z$, likely determines $Y$ and correlates with $M$, $X$, or both.[1] When $Z$ meets these conditions, failing to control for $Z$ will produce omitted variable bias (OVB) in one or more of the coefficient estimates, so researchers expand (1) as follows:

$$Y = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 X \times M + \beta_4 Z + \varepsilon \quad (2)$$

Researchers are generally diligent about identifying and controlling for $Z$. However, avoiding OVB when testing interactive effects (e.g., $\beta_3$) often requires interacting $Z$ with $X$ or $M$

[1] For simplicity, we base our discussion on regressions with one $Z$, but the same issues relate to settings with multiple $Z$s.

(i.e., *X*×*Z* or *M*×*Z*). The importance of including these interactive controls is often overlooked. The objective of this study is to provide plain-English guidance of when, why, and how to use interactive controls in when testing moderating effects.

We discuss the role of interactive controls and how to interpret coefficient estimates in the context of generic accounting scenarios using simple simulations. We vary the scenario by designating explanatory variables as either *exogenous* or *endogenous*. We define an *exogenous variable* as an explanatory variable that is uncorrelated with the regression error term (i.e., unrelated to unmodelled factors that determine *Y*). Likewise, an *endogenous variable* is correlated with the regression error term (i.e., is related to unmodelled factors that determine *Y*).[2,3]

Our first scenario assumes *X* and *M* are exogenous. In this setting, unbiased estimation of the effects of *X* or *M* does not require a control for *Z* because *Z* does not correlate with *X* or *M*. Likewise, unbiased estimation of the interaction *X*×*M* does not require control for interactions *X*×*Z* or *M*×*Z* because, conditional upon *X* and *M* being exogenous, neither *X*×*Z* or *M*×*Z* can correlate with *X*×*M*. We use this scenario to develop intuition for how the interaction term *X*×*M* affects coefficient estimates and interpretations.

Next, we consider a scenario with an exogenous *X* and an endogenous *M*. When *M* is endogenous, researchers need to identify factors (*Z*) that correlate with *M* (but are not outcomes

---

[2] The technical definition of *endogeneity* is that the expected regression error conditioned on a set of variables is not equal to zero; i.e., $E(\varepsilon|X) \neq 0$. As discussed by Wooldridge (2013, p. 87), the term *endogenous variable* is commonly used to describe a variable that is correlated with the error term. The concept of endogeneity is necessarily model-specific, so a given variable can be endogenous in a model without controls but conditionally exogenous with the inclusion of appropriate controls. This paper specifically focuses on endogeneity caused by a correlated omitted variable, but endogeneity can also arise from measurement error or simultaneity.
[3] While we categorize variables as either endogenous or exogenous, in practice, this distinction falls on a continuum. For example, in accounting settings, some variables are more endogenous than others and nearly all non-experimental variables are endogenous to a certain degree. In causal inference, the degree of endogeneity will determine the extent of bias in estimates of the true causal effect. While our primary simulations assume at least one variable (*X* or *M*) is exogenous, Section 5.2 discusses challenges faced when this assumption is relaxed.

of *M*) and that also moderate the relation between *X* and *Y*. Said differently, when using $X \times M$ to test the effects of a specific moderator *M*, other moderators of the relation between *X* and *Y* that also correlate with *M* need to be controlled. This requires controlling for both *Z* and the interaction between *Z* and *X* (i.e., $X \times Z$). For the same reason, the reverse is true when *X* is endogenous and *M* is exogenous. That is, obtaining an unbiased estimate of the effect of $X \times M$ requires controlling for $Z \times M$. The key intuition is that *Z* should include controls that correlate with whichever of *X* and *M* is endogenous, and these controls generally need to be interacted with whichever of *X* and *M* is exogenous.[4,5]

A summary of best practices in tests of interactive effects is provided in an attached Appendix. That said, we caution that we do not provide a "one-size fits all" strategy for proper control in interactive settings. Designing well-specified tests is context-specific and requires making informed assumptions about things we cannot observe or test.

## 2. Scenario 1: All variables are exogenous

### *2.1 Intuition*

For our first scenario, we designate both our *X* and *M* as exogenous variables. We use a common setting that frequently involves interactive effects and where variables are often treated as exogenous, short-window "earnings response coefficients" or "ERCs." Throughout this study and unless otherwise noted, we assume that all models are correctly specified, and all variables are measured without error. We begin our discussion by considering a regression of abnormal

[4] For simplicity, we refer to *Z* as "correlating" with the test variable rather than "causing" the test variable. It is important to note that correlating variables may be "bad controls" in the context of causal inference. See Whited et al. (2022) for background on appropriate controls. In this study, we assume that correlating controls meet the criteria of "good controls" discussed in that paper.
[5] The issues discussed in this paper apply more generally to tests of any "interactive" effects. For example, researchers often perform a difference-in-differences analysis with a treatment indicator (*Treat*) interacted with a post- indicator (*Post*). In accounting research, *Treat* is often endogenous (correlated with *Z* variables). As our paper illustrates, the setting would also require the inclusion of $Z \times Post$ as controls.

stock returns ($Y$) on earnings surprise ($X$):

$$CAR = \alpha_0 + \alpha_1 UE + \varepsilon \quad (3)$$

*CAR* is the cumulative abnormal stock return immediately around an earnings announcement and *UE* is a measure of unexpected earnings contained in the earnings announcement. The regression intercept, $\alpha_0$, represents the average *CAR* when *UE* equals zero. $\alpha_1$ is the ERC, which represents the average change in *CAR* for a one-unit increase in *UE*. Intuition and prior research suggest that $\alpha_1$ should be positive.

ERC research frequently considers how the ERC varies with some other factor, which we refer to as a moderating variable ($M$). For example, suppose we want to know whether firm size (*Size*) moderates the relation between earnings surprises and returns. We can test this relation by adding an interaction to equation (3):

$$CAR = \beta_0 + \beta_1 UE + \beta_2 Size + \beta_3 UE \times Size + \varepsilon \quad (4)$$

We change the coefficient notation from $\alpha$ in (3) to $\beta$ in (4) to emphasize that the two models investigate different questions: (3) estimates the average effect of *UE* on *CAR* across the full sample, while (4) estimates the joint effects of *UE* and *Size* on *CAR*. The interaction term in (4), $UE \times Size$, allows the ERC to depend on the value of *Size*. Thus, for any given observation $i$, the estimated ERC would equal ($\beta_1 + \beta_3 \times Size_i$). $\beta_1$ is the ERC when *Size* equals zero (i.e., $\beta_1 + \beta_3 \times 0 = \beta_1$), and the interaction coefficient $\beta_3$ indicates how the ERC incrementally changes with a one-unit increase in *Size*. Similarly, $\beta_2$ is the effect of *Size* on *CAR* when *UE* equals zero.

If *UE* and *Size* are uncorrelated with all other determinants of *CAR* in (4), *UE* and *Size* do not correlate with the error term and are therefore exogenous. Thus, estimating (4) will produce unbiased coefficient estimates.

### *2.2 Simulation*

To illustrate the concepts above, we simulate data on earnings, firm size, and returns using the following data generating process (DGP). We provide full Stata code to replicate these analyses in an online appendix:

Step 1: Create 10,000 observations.

Step 2: Generate an unexpected earnings (*UE*) variable as a random draw from a normal distribution with a mean of 0 and a standard deviation of 2. The mean of 0 indicates that the market's earnings expectation is correct on average.[6]

Step 3: Generate a company size (*Size*) variable as a random draw from a normal distribution with a mean of 8 and a standard deviation of 2. The positive mean is consistent with firms having positive market values.[7] Note that *Size* is uncorrelated with *UE*, meaning that bigger firms do not have systematically more positive or negative UE.

Step 4: Generate cumulative abnormal returns (*CAR*) using (4) and the following parameters: $\beta_0 = 0$; $\beta_1 = 10$; $\beta_2 = 0$; $\beta_3 = 10$; and $\varepsilon$ is randomly drawn from a normal distribution with a mean of 0 and a standard deviation of 100. Substituting these parameters into (4) yields the following:

$$CAR = 0 + 10 \times UE + 0 \times Size + 10 \times UE \times Size + \varepsilon \quad (5)$$

Equation (5) establishes $\beta_1$ from (4) as equal to 10 and indicates that *CAR* increases by 10 units for each one-unit increase in *UE* when *Size* equals zero. Similarly, equation (5) sets $\beta_3$ from (4) equal to 10 and indicates that the relation between *CAR* and *UE* increases by 10 units for each

[6] Using analyst forecasts for "expected" earnings, studies often document positive average earnings surprises. This occurs due to the on-average pessimistic bias in analyst forecasts. We specify a zero mean for simplicity and interpretation, but we can use a non-zero mean for *UE* without loss of generality.

[7] Below we discuss the case when *Size* has a zero mean (i.e., when *Size* is demeaned), in which case the uninteracted terms in a regression have different coefficient estimates and interpretations. However, demeaning a variable has no bearing on the presence or absence of omitted variable bias.

one-unit increase in *Size*. $\beta_2$ of zero indicates that firm size does not, on average, relate to *CAR*. Instead, *Size* only affects *CAR* by moderating the relation between *UE* and *CAR*.

We iterate the DGP above 1,000 times and report average outputs from a series of regressions (coefficients, t-stats, and adjusted $R^2$) in Table 1. We omit intercept estimates for brevity.

We first estimate (3) with no moderating effects to illustrate the average ERC in the sample. As shown in column (i), the ERC estimate ($\widehat{\alpha_1}$) in this specification is approximately 90.[8] This does not equal the parameter on *UE* in (5), $\beta_1$ of 10, because $\widehat{\alpha_1}$ estimates the average ERC across all observations, whereas $\beta_1$ in (5) represents the ERC conditional on *Size* set equal to 0 (i.e., due to the inclusion of the term *UE*×*Size*). If we substitute the sample mean of *Size* (which we specified to be eight) into (5), we see that $\widehat{\alpha_1}$ approximates the sample average ERC as specified in the DGP: $10 \times UE + 10 \times UE \times 8 = 90 \times UE$.

Column (ii) presents estimates of (3) after adding *Size* but not the interaction (*UE*×*Size*) to the model. In this specification, $\widehat{\alpha_1}$ reflects an estimate of the average relation between *UE* and *CAR* after controlling for *Size,* but it does not inform whether larger firms have larger or smaller ERCs. $\widehat{\alpha_1}$ in column (ii) does not change from column (i) because the DGP specified *Size* to be uncorrelated with *UE*, so "the effect of *UE* on *CAR* holding *Size* constant" does not differ from "the effect of *UE* on *CAR*."

Column (iii) presents estimates of (4) which considers the moderating effect of *Size* on the ERC. $\widehat{\beta_3}$ in column (iii) of Table 1 indicates that the relation between *UE* and *CAR* increases by 10 for each one-unit increase in *Size*. We can estimate the relation between *UE* and *CAR* for a

---

[8] We discuss estimates using a hat (e.g., $\widehat{\alpha}$ or $\widehat{\beta}$) to indicate that they reflect estimates, rather than the actual parameters defined in the DGP.

particular firm *i* as (9.98 + 10.00×$Size_i$). Thus, column (iii) estimates that the average firm (*Size* equal to 8) has an ERC of (9.98+10.00×8) = 89.98, which (uncoincidentally) approximates the average ERC estimated in column (i).

A key takeaway from Table 1 is that including *UE*×*Size* changes the interpretation of the coefficient on *UE* in column (iii) (i.e., $\widehat{\beta_1}$) relative to in column (i) (i.e., $\widehat{\alpha_1}$). In column (i), $\widehat{\alpha_1}$ reflects the average ERC across all firms; in column (iii), $\widehat{\beta_1}$ reflects the expected ERC for firms with *Size* = 0. The difference in the magnitude and significance of these two estimates illustrates the importance of interpreting models in light of included interactions. For example, if we mistakenly interpreted $\widehat{\beta_1}$ from column (iii) as the sample average effect of *UE* on *CAR*, we would infer that *UE* has a much more modest effect on *CAR* relative to *UE*'s true average effect (accurately estimated in column (i)). Moreover, $\widehat{\beta_1}$ in column (iii) is conditional on *Size* = 0, so it is unlikely to be particularly interesting given that few firms have *Size* of zero.[9]

***2.3 Interpreting economic magnitudes***

Because coefficients on the stand-alone terms (or "main effects") in interaction models can have uninteresting interpretations unto themselves, it is often useful to demean all independent variables before interacting variables and estimating the regression coefficients (Burks et al. 2019). Demeaning "centers" variables around the sample average so that stand-alone terms estimate meaningful effects at the sample average of the interacted variables.[10] To illustrate, we demean *Size* (the mean of *UE* is already 0), recompute the interaction *UE*×*Size,* and

[9] We draw *Size* from a normal distribution with a mean of 8 and a standard deviation of 2, so a 0 (or negative) value for firm size is technically possible though highly unlikely in our simulation (a 4σ event – which is expected to occur approximately 1:30,000 observations). In practice, though, few firms should have a size of zero by most common metrics (e.g., assets, market value of equity). It is not uncommon to find counterintuitive coefficient estimates on stand-alone variables when interactions condition the "main effect" on uninteresting values of a moderating variable.

[10] "Demeaning" involves calculating the sample average of each variable, and then subtracting that average from each observation. "Standardizing" goes a step further and divides the demeaned variable by its standard deviation, such that the resulting variable is in units of standard deviations.

re-estimate equation (4). We present results in column (iv) of Table 1.

De-meaning does not affect interaction coefficient estimates, so the estimate on (demeaned) *UE*×*Size* in column (iv) is identical to the estimate on the interaction in column (iii) where *Size* is not demeaned. However, $\widehat{\beta_1}$ differs from column (iii) because $\widehat{\beta_1}$ in column (iv) estimates the ERC for a firm with (demeaned) *Size* equal to zero, which is by design the sample average firm size (8 per our DGP). Thus, $\widehat{\beta_1}$ equals 89.99 in column (iv), which is the same as the average ERC estimated in column (i), $\widehat{\alpha_1}$. In short, demeaning the interacted variables often yields more easily interpretable stand-alone estimates.[11]

Demeaning can also assist with interpreting the economic magnitude of moderating effects. To gauge the economic magnitude of a moderating effect like in column (iii), researchers sometimes compare the magnitudes of the coefficients on the interactive term relative to the stand-alone effect. For example, a researcher may make a relative statement such as "a one-unit increase in *Size* in column (iii) drives a 10.00/9.98 ≈ 100% increase in ERCs," which would appear to be quite a large effect. However, because $\widehat{\beta_1}$ in (iii) reflects the estimated ERC at *Size = 0*, this relative comparison only holds true for a one-unit increase from a baseline of *Size* = 0. Researchers should instead carefully select reasonable values of moderating variables for making relative comparisons. In our setting, a more meaningful interpretation would be to estimate the effect of a one-unit increase in *Size* on the ERC for a firm of average size (i.e., *Size = 8)*. To do so, we would use the estimated effect on the ERC of a one-unit increase in *Size* of 10.00 (from column (iii)) and the ERC for a firm of average size of 89.98 (9.98 + 10.00×8 = 89.98) to calculate the economic magnitude of a one-unit increase for an average-sized firm of 11%

[11] Demeaning indicator variables can sometimes yield less meaningful stand-alone effects, so researchers should thoughtfully consider variable specification before demeaning. With indicator variables in particular, a value of 0 may reflect a meaningful and realistic value. If all variables are not demeaned, though, then the "main effect" estimate will not be at the sample average of all variables.

(10.00/89.98). By demeaning *Size* in (iv), we can just divide the coefficient on *UE*×*Size* by the coefficient on *Size* (10.00 ÷ 89.99) to quantify the estimated moderating effect for an averaged size firm.

***2.4 Collinearity between regressors***

As also discussed in Burks et al. (2019), researchers sometimes express concerns about high collinearity between regressors in models with interacted controls. For example, in column (iii) of Table 1, *UE* and *UE*×*Size* are correlated at 97% and variance inflation factors (VIFs), a common regression diagnostic for gauging collinearity, are approximately 17 for *UE* and *UE*×*Size*. A researcher might therefore conclude that it is impractical to include *UE* and *UE*×*Size* in the same regression due to collinearity. However, comparing column (iii) to column (iv) elucidates why using VIFs to diagnose collinearity concerns is problematic, particularly when high VIFs relate to interaction terms. In column (iii), *Size* has a positive mean, leading to high correlations between the terms *UE* and *UE*×*Size* by construction. The VIF for *UE*×*Size* is then calculated as $1/(1-R^2)$, where $R^2$ is from a regression of *UE*×*Size* on *UE* and *Size*. The high correlation between *UE* and *UE*×*Size* leads to a high $R^2$ and VIF. By comparison, in column (iv), de-meaning *Size* leads to a correlation between *UE* and *UE*×*Size* of approximately 0, decreasing the VIF to 1, its minimum possible value. Thus, although columns (iii) and (iv) yield identical statistical conclusions (identical coefficient and t-stat on the interaction and identical $R^2$), relying on VIFs to "diagnose collinearity" would lead to the erroneous conclusion that collinearity is concerning in column (iii) but not in column (iv). This simple simulation highlights that VIFs are not a reliable approach to assess collinearity in interactive settings.

## 3. Scenario 2: exogenous *X* and endogenous *M*

Our first scenario assumes that both *X* (*UE*) and *M* (*Size*) are unrelated to other factors affecting *Y* (i.e., they are exogenous). In practice, however, researchers often wish to test interactive effects where *M* is at least partially determined by other *Z* variables (i.e., *M* is endogenous). Identifying the moderating effect of *M* in these settings may require disentangling the effects of the interaction $X \times M$ from the effects of interactions between *X* and *Z* variables. The next subsection illustrates this point by expanding upon our earlier ERC simulation.

### *3.1 Simulation*

This scenario uses the same DGP as before, in which *CAR* is a function of *UE* and *Size* as specified in (5). Our only change is to create a new indicator variable, *WSJ*, which equals 1 if the firm's earnings announcement has an associated *Wall Street Journal* media article. The media tend to write articles about large firms, so we construct *WSJ* to be a positive function of *Size*. For simplicity, we assume that the media always writes stories about the largest 50 percent of firms, and never write stories about the smallest 50 percent, which adds the following step to our simulation:[12]

Step 5: Generate a *Wall Street Journal* coverage variable (*WSJ*) which equals 1 for firms above the median for *Size* and 0 otherwise.

Suppose that a researcher hypothesizes that media coverage of a firm's earnings announcement draws attention to the firm and causes a larger price response per unit of earnings surprise, which should manifest as larger ERCs. In this setting, *WSJ* is the moderating variable (*M*) in a regression of *CAR* (*Y*) on *UE* (*X*). The researcher plans to use the interaction $UE \times WSJ$ to test their hypothesis using the following model:

$$CAR = \beta_0 + \beta_1 UE + \beta_4 WSJ + \beta_5 UE \times WSJ + \varepsilon \tag{6}$$

Note that we use $\beta_4$ and $\beta_5$ because neither *WSJ* nor $UE \times WSJ$ appear in our DGP in (5).

[12] The specific function by which *Size* affects *WSJ* is unimportant. The inferences relating to OVB hold for any non-zero correlation between firm size and media coverage.

Because *WSJ* and *UE*×*WSJ* do not appear in the DGP, media coverage has no effect on *CAR,* either directly or by moderating the ERC. In other words, unbiased estimates of both $\beta_4$ and $\beta_5$ will equal 0.

Because *WSJ* and *Size* are correlated, failing to consider *Size* in (6) will produce biased coefficient estimates. As a rule, if any two variables are correlated (*WSJ* and *Size* in this setting), then interactions between those variables and any random variable (*UE* in this setting) will also be correlated. Thus, in our simulation, *UE*×*WSJ* and *UE*×*Size* are necessarily correlated.[13] *UE*×*Size* is also a determinant of *Y,* so omitting *UE*×*Size* from (6) will produce a biased estimate of the moderating effect of *WSJ* on the ERC, $\widehat{\beta_5}$ .

Column (i) of Table 2 presents the output from (6) using our simulated data and demeaned versions of our variables as recommended in Section 2. The significantly positive coefficient estimate on *UE*×*WSJ*, $\widehat{\beta_5}$, suggests that media coverage increases ERCs. However, because *WSJ* does not appear in the DGP in (5), we know that this is an erroneous inference.

A researcher concerned about OVB in column (i) may attempt to address this bias by controlling for *Size* as follows:

$$CAR = \beta_0 + \beta_1 UE + \beta_2 Size + \beta_4 WSJ + \beta_5 UE \times WSJ + \varepsilon \tag{7}$$

However, as shown in Table 2 column (ii), controlling for *Size* has no material impact on $\widehat{\beta_5}$. This occurs because even after controlling for *Size*, *UE*×*WSJ* is still correlated with *UE*×*Size,* which column (ii) omits.

Adding the control for *UE*×*Size* results in the following regression:

[13] The correlation between the interactions is simply the covariance scaled by the sum of individual variances. To illustrate, refer to the DGP used for the simulation in this section. It can be shown that: Cov(*Size*×*UE*, *WSJ*×*UE*) = Cov(*Size*, *WSJ*)×Var(*UE*) + E(*Size*)×E(*WSJ*)×Var(*UE*) + E(*UE*)$^2$×Cov(*Size*, *WSJ*). Assuming *Size* is demeaned, E(*Size*) and E(*UE*) are both zero, so this expression simplifies to the first term (Cov(*Size*, *WSJ*)×Var(*UE*)). The simulation parameters specify that *Cov*(*Size*, *WSJ*) > 0.

$$CAR = \beta_0 + \beta_1 UE + \beta_2 Size + \beta_3 UE\times Size + \beta_4 WSJ + \beta_5 UE\times WSJ + \varepsilon \quad (8)$$

This specification produces unbiased estimates consistent with the DGP in (5), which we present in column (iii).[14] Neither $\widehat{\beta_4}$ nor $\widehat{\beta_5}$ is significant once properly controlling for *UE*×*Size,* and $\widehat{\beta_1}$, $\widehat{\beta_2}$, and $\widehat{\beta_3}$ approximate the parameters specified by the DGP.

### *3.2 Takeaways with an exogenous X and endogenous M*

ERCs are a common setting in which we encounter a (largely) exogenous treatment, *X*, and an endogenous moderating variable, *M*. "Natural experiments," or designs where a subset of observations is subject to as-if random treatment assignment (e.g., the Reg SHO pilot study), also fit this description if a researcher wants to test whether the effects of the natural experiment vary with an endogenous firm characteristic (*M*).

In settings with an exogenous *X* and endogenous *M*, we suggest first identifying control variables (*Z*) that correlate with *M* (but, as usual, are not outcomes of *M*). If *Z* also moderates the relation between *X* and *Y* (a common occurrence), the interaction term, *X*×*Z*, is needed to avoid omitted variable bias. In other words, when evaluating how an endogenous variable, *M*, impacts the "slope" of interest (e.g., $\partial CAR/\partial UE$), researchers should identify and control for the effects of other factors, *Z*, that are both correlated with *M* and may also moderate the relation between *X* and *Y* by including the vector of interactive controls, *X*×*Z.*

## 4. Scenario 3: endogenous *X* and an exogenous *M*

While Scenario 2 examines an exogenous *X* (*UE*) and endogenous *M* (*WSJ*), similar considerations apply to the scenario in which *X* is endogenous and *M* is exogenous. For example, studies often put forth a primary prediction of how an endogenous *X* affects *Y*, and then follow

[14] Recall that we use demeaned firm size in these specifications, so the coefficient estimates conform to the DGP (similar to column (iv) from Table 1).

with a secondary test of how an exogenous *M* moderates that relation. While truly exogenous *M* are difficult to find, recent studies attempt to leverage plausibly exogenous *M* stemming from natural experiments.

In these settings, researchers are usually diligent about identifying *Z* variables that correlate with *X* and may also affect *Y* (i.e., confounders), and controlling for these *Z* in their first set of tests. It is also critical, though, to consider whether the exogenous *M* also moderates the relation between *Z* and *Y*. Because *X* and *Z* are correlated, the interactions $X \times M$ and $Z \times M$ are also correlated. Thus, if *M* moderates the relation between *Y* and *Z*, then $Z \times M$ is a correlated determinant of *Y* that must be controlled as follows:

$$Y = \alpha + \beta_1 X + \beta_2 M + \beta_3 X \times M + \beta_4 Z + \beta_5 Z \times M + \varepsilon \qquad (9)$$

The intuition for this scenario is largely the reciprocal of our previous scenario, so for brevity we do not provide a simulation. The takeaway is that if either *X* or *M* are endogenous, then the interaction between *X* and *M* will also be endogenous. When testing interactive effects in these scenarios, researchers should take care to control for the interactions between the endogenous variable and the controls.

## 5. Other considerations

### *5.1 Sample partitions on M*

Accounting research frequently includes a long list of regressors, in which case it can be cumbersome to interact each variable with *M* and to clearly present results. When *M* is dichotomous and exogenous, partitioning the sample on *M* can serve the same econometric purpose as interacting *M* with all regressors.[15, 16] Partitioning involves estimating a regression

[15] In models with fixed effects, regressions with interactions versus partitioned regressions are only equivalent when the fixed effects are also interacted with *M*.
[16] We discuss partitions with respect to *M*. However, if *X* is dichotomous and exogenous, partitioning on *X* can effectively test the interactive effect $X \times M$.

model separately on subsamples of *M* equals zero and *M* equals one. Comparing the coefficient estimates on *X* between the partitioned regressions is equivalent to testing the moderating effect of *M* on *X* with the interaction *X*×*M*.

Partitioning the sample based on an endogenous *M* is less effective because doing so effectively partitions on the correlated variable(s) as well. For example, consider testing the moderating effect of analyst following on the relation between *UE* and *CAR* by partitioning on whether the firm is followed by at least one analyst (*Analyst*). Because *Analyst* will be strongly positively correlated with firm size, the firms in *Analyst=1* sample will be larger on average than the firms in the *Analyst=0* sample. In this case, the difference in coefficient estimates on *UE* between partitions may reflect the moderating effects of firm size rather than analyst following.

### *5.2 Interactive controls when both X and M are endogenous*

Throughout the paper, we have assumed either *X* or *M* is exogenous. However, studies often interact two endogenous variables, which further complicates effective control in interactive models. For example, a control variable, *Z,* from the primary test of the effects of *X* on *Y* then serves as the moderating variable, *M*, in a subsequent test. If *Z* is endogenous, then the interaction of interest will be of two endogenous variables.

Identifying proper controls for one construct (i.e., either *X* or *M*) is challenging, so ensuring *Z* includes controls necessary for *both X* and *M* is even more so. Practically speaking, it is likely impossible to fully address OVB in the presence of two interacted, endogenous variables. If the variables are only weakly endogenous, then interacting *Z* with both *X* and *M* may allow some reasonable inference from *X*×*M*.[17] However, stronger correlations likely imply

[17] By weakly endogenous, we mean that the variables are weakly correlated with observable and presumably unobservable traits that may also influence *Y*. Exogeneity (or weak endogeneity) can only be justified theoretically; however, if the variables exhibit very low correlations with observable variables in the study, this may suggest that endogeneity concerns are less important.

greater endogeneity concerns and will be more difficult to control. Furthermore, if there is more than one *Z* variable, each combination of $Z_1 \times Z_2$ likely also correlates with $X \times M$, so $Z_1 \times Z_2$ interactions are also potential confounding variables.

Overall, we urge researchers to use caution when drawing strong causal inferences from models that interact more than one endogenous variable. When interacting two endogenous variables is necessary, researchers should clearly recognize the complexities of the setting and the associated limitations of interpreting such models.

## 6. Concluding thoughts and suggestions

This study discusses the conditions under which interactive test variables require interactive controls for unbiased estimation of moderating effects and provides simulations to illustrate concepts in common accounting contexts. For reference, we summarize our recommendations for interactive controls in an Appendix. In short, when testing the interaction between an endogenous and an exogenous variable, researchers should include interactions between the exogenous variable and control variables, *Z*, that correlate with the endogenous variable (either *X* or *M*). In general, we urge caution when testing the interaction between two endogenous variables.

We hope that this study provides practical insights and guidance for researchers testing moderating effects. Especially interested readers can refer to Whited et al. (2022) for a discussion of the characteristics of “good” versus “bad” controls; Jollineau and Bowen (2021) for further discussion of moderating versus mediating effects, Burks et al. (2019) for a more detailed investigation into interaction effects; and Breuer and deHaan (2022) for discussion of control variables in the form of fixed effects.

**References**


Burks, J., D. Randolph, and J. Seida, 2019. Modeling and interpreting regressions with interactions. *Journal of Accounting Literature* 42: 61-79.

Breuer, M. and deHaan, E. 2022. Using and interpreting fixed effects models. *SSRN working paper.*

Jollineau, J., and R. Bowen. 2022. A practical guide to using path analysis: mediation and moderation in accounting research. *Journal of Financial Reporting*, forthcoming.

Wooldridge, J. M. 2013. Introductory Econometrics: A Modern Approach 5th Edition

Whited, R. L., Q. T. Swanquist, J. E. Shipman, and J. R. Moon. 2022. Out of control: The (over)use of controls in accounting research. *The Accounting Revie*w 97 (3): 395-413.

## Table 1: Basics of Stand-alone Effects and Interaction Effects

This table presents the average coefficient estimates (in bold), t-statistics (in italics), and adjusted r-squares from regressions based on 1,000 simulated datasets following the DGP detailed in Section 2 and in the code in our online appendix. *Size* is not modified in columns (ii)-(iii) and is demeaned in column (iv). *** indicates statistical significance at a 1 percent level of confidence or better.

| | | (i) | (ii) | | (iii) | (iv) |
|---|---|---|---|---|---|---|
| | | *CAR* | *CAR* | | *CAR* | *CAR* |
| *UE* | $\alpha_1$ | **89.99*****| **89.99***** | $\beta_1$ | **9.98***** | **89.99***** |
| *t-stat* | | *167.11* | *167.11* | | *4.84* | *179.95* |
| *Size* | $\alpha_2$ | | **0.00** | $\beta_2$ | **0.00** | **0.00** |
| *t-stat* | | | *0.01* | | *0.01* | *0.01* |
| *UE×Size* | | | | $\beta_3$ | **10.00***** | **10.00***** |
| *t-stat* | | | | | *39.98* | *39.98* |
| Adjusted $R^2$ | | 0.736 | 0.736 | | 0.773 | 0.773 |
| N | | 10,000 | 10,000 | | 10,000 | 10,000 |

## Table 2: Interaction Controls with Endogenous *M*

This table presents the average coefficient estimates (in bold), t-statistics (in italics), and adjusted r-squares from regressions based on 1,000 simulated datasets following the DGP detailed in Sections 2-3 and in the code in our online appendix. *Size* is demeaned in all regressions. *** indicates statistical significance at a 1 percent level of confidence or better.

| | | (i)<br>*CAR* | (ii)<br>*CAR* | (iii)<br>*CAR* |
|---|---|---|---|---|
| *UE* | $\beta_1$ | **74.02*** ** | **74.02*** ** | **89.94*** ** |
| *t-stat* | | *101.78* | *101.77* | *92.84* |
| *Size* | $\beta_2$ | | **0.03** | **0.03** |
| *t-stat* | | | *0.04* | *0.04* |
| *UE×Size* | $\beta_3$ | | | **9.98*** ** |
| *t-stat* | | | | *24.04* |
| *WSJ* | $\beta_4$ | **-0.05** | **-0.15** | **-0.14** |
| *t-stat* | | *-0.02* | *-0.05* | *-0.04* |
| *UE×WSJ* | $\beta_5$ | **31.95*** ** | **31.95*** ** | **0.11** |
| *t-stat* | | *31.06* | *31.06* | *0.07* |
| Adjusted R-squared | | 0.760 | 0.760 | 0.773 |
| N | | 10,000 | 10,000 | 10,000 |

## Appendix: Interactive Controls Best Practices Summary

This appendix summarizes general best practices for including interaction terms in regression tests. Note, these are not "one-size fits all" recommendations. Researchers should carefully evaluate endogeneity threats and research design choices within the context of each setting.

| | **Exogenous Moderator (*M*)** | **Endogenous Moderator (*M*)** |
|---|---|---|
| **Exogenous Treatment (*X*)** | No controls are needed for unbiased estimation of the effects of $X$ or $X \times M$ on $Y$.<br><br>As is true for all four quadrants in this table, including controls that explain $Y$, but are not outcomes of $X$ or $M$, can improve model fit and reduce standard errors. However, these variables are not necessary for unbiased estimation. | Identify variables, $Z$, that correlate with $M$ and determine $Y$. If $Z$ requires control for unbiased estimation, then also control for the interactions $X \times Z$. |
| **Endogenous Treatment (*X*)** | Identify variables, $Z$, that correlate with $X$ and determine $Y$. If $Z$ requires control for unbiased estimation, then also control for the interactions $M \times Z$. | Identify variables, $Z$, that correlate with $X$ and/or $M$ and determine $Y$. Control for $Z$, $M \times Z$, $X \times Z$, and possibly combinations of $Z_1 \times Z_2$.<br><br>However, these may still not fully address OVB. In general, researchers should recognize the limitations of testing the interaction between two endogenous variables. |