A Convenient Omitted Variable Bias Formula for Treatment Effect Models

Damian Clarke

Universidad de Santiago de Chile

10 March 2018

A Convenient Omitted Variable Bias Formula for Treatment Effect Models*

Damian Clarke

March 10, 2018

Abstract

Generally, determining the size and magnitude of the omitted variable bias (OVB) in regression models is challenging when multiple included and omitted variables are present. Here, I describe a convenient OVB formula for treatment effect models with potentially many included and omitted variables. I show that in these circumstances it is simple to infer the direction, and potentially the magnitude, of the bias. In a simple setting, this OVB is based on mutually exclusive binary variables, however I provide an extension which loosens the need for mutual exclusivity of variables, and derives the bias in difference-in-differences style models with an arbitrary number of included and excluded "treatment" indicators.

JEL codes: C21, C22, C13.
Keywords: Omitted variable bias; Ordinary Least Squares Regression; Treatment Effects; Difference-in-Differences.

*This research was supported by FONDECYT (grant number 11160200) of the Government of Chile. I thank Serafima Chirkova for useful comments. Affiliation: Department of Economics, Universidad de Santiago de Chile. Contact email: damian.clarke@usach.cl.

1 Introduction

The omitted variable bias (OVB) is a staple of econometrics courses, and applied research across all fields of economics, appearing as early as Theil (1957). In its most basic form, the omission of a single relevant explanatory variable in a linear model leads to an elegant bias formula providing a simple link between parameter estimates, true values, and underlying relationships between variables. This formula is often amenable to analysis using intuition from economic models. However, once outside a simple text-book case, the omitted variable bias can become increasingly complex, such that inferring even the direction of bias is impossible.[1] As laid out in Clarke (2005), this has the extreme consequence where the inclusion of a greater or smaller number of a partial set of the full controls in a model has an indeterminate effect on the bias.

In this paper I provide a convenient representation of the omitted variable bias in a model with arbitrarily many included and omitted variables, with an intuitive link to the underlying economic process described by the model. The convenience of this representation comes at the cost of the *class* of models for which it serves. This representation is provided for models based on a series of mutually exclusive binary "treatment" variables. After documenting the bias for a case where potentially many treatment variables are included and excluded in a linear model, I then provide an extension to a more complicated treatment effect model: the difference-in-differences model. In this setting, while treatment effect indicators may be mutually exclusive among themselves, common fixed effects (for example for time) are shared. I document in this case that the simple OVB formula holds, and follows the same logic as in the simple static models.

While this is a restrictive model, it is nonetheless frequently observed in empirical applications. Models designed to estimate treatment effects with multiple treatment statuses—where a population is split into various treatment groups and a control group—are often encountered. A number of such cases are found in (Kremer, 2003; Banerjee et al., 2007), (and indeed, these are referred to as "cross-cutting designs" in Duflo et al. (2007)) and also in areas outside of economics, for example in medicine (Baron et al., 2013). This is particularly relevant in cases where concerns exist about imperfect observation of treatment status, for example where treatment externalities occur (Miguel and Kremer, 2004), or where environmental or other shocks may diffuse over space (Almond et al., 2009) complicating the precise definition of treatment. Finally, the OVB derived here extends to *any* model based on mutually exclusive (or largely mutually exclusive) binary independent variables, such as models

[1] For example, consider the presentation of Greene (2002, p. 180) in his widely studied text-book, who states, "if more than one [omitted] variable is included, then the terms in the omitted variable formula involve multiple regression coefficients, which themselves have the signs of partial, not simple, correlations. …This requirement might not be obvious, and it would become even less so as more regressors were added to the equation."

with a large number of fixed effects.

In what follows, I very briefly describe the traditional omitted variable bias, before describing the simplified convenient OVB formula for treatment effects models. I then document that this bias can be generalised to more complicated cases, namely difference-in-differences models, without losing its simple interpretation.

2 The Traditional Omitted Variable Bias in Linear Models

To illustrate the traditional omitted variable bias model, consider a correctly specified regression model of the form:

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \tag{1}$$

Here the independent variables $X = [X_1 \quad X_2]$ are split into two matrices. The X_1 matrix consists of $1 + k_1$ variables (a constant plus k_1 other independent variables) and X_2 consists of k_2 variables. The stochastic error term ε in the full model is orthogonal to each column of X in 1. In what remains of this paper, in the case that we are working with particular variables, rather than matrices, these are denoted in lower-case letters. Thus, X_1 refers to the full set of included variables (which includes the constant terms), while x_1 refers to a single included variable, called x_1.

If the dependent variable y is regressed *only* on X_1 the expectation of the OLS estimator of the parameter β_1 is:

$$E[\widehat{\beta}_1^{ovb}|X] = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2. \tag{2}$$

This is the well-known omitted variable bias formula. Here the bias term of $(X_1'X_1)^{-1}X_1'X_2\beta_2$ is sometimes written as $\delta\beta_2$, where $\delta \equiv (X_1'X_1)^{-1}X_1'X_2$ is a $(1 + k_1) \times k_2$ matrix from the regression of each variable in X_2 on the full set of X_1 variables. This "textbook" omitted variable bias has a very simple interpretation in the case where $k_1 = 1$ and $k_2 = 1$: it is the product of the simple correlation between X_1 and X_2 (δ) and the direct effect of X_2 on y in the data generating process (β_2). However, this formula can quickly become unwieldy when $k_1 > 1$ as each element of the $\widehat{\beta}_1^{ovb}$ vector differs from β_1 in expectation depending on the partial correlation between the relevant variable in X_1, conditional on the remainder of X_1. Economic models and intuition often have less to say about *partial* correlations between variables, especially when conditioning on many relevant variables.

3 A Convenient Omitted Variable Bias Formula for Mutually Exclusive Variables

Consider now model 1, where each element of X is a binary variable. This is what we refer to here as a treatment effect model which examines the impact of multiple "treatments" on an outcome of interest. Further, assume that each variable is mutually exclusive. Such a model is common in cases where a pool of subjects are split into various treatment groups and a control group, each receiving at most one treatment. While this mutual exclusivity assumption may appear limiting, it is actually more flexible than it may first appear, given that receipt of multiple treatments can be considered as a treatment unto itself.[2] In this case, we can show that the omitted variable bias above has a very convenient and intuitive form, and is easily linked to economic theory, even in cases where an arbitrary number of included and omitted variables exist. At the end of this section, we then loosen the assumption that each variable is mutually exclusive.

3.1 A Single Omitted Variable and Included Variable

In the most simple case of a single omitted variable (an $N \times 1$ vector x_2) and single included explanatory variable (an $N \times 1$ vector x_1), plus an intercept term, the omitted variable bias is easily interpretable as per equation 2. Nonetheless, I briefly document an alternative interpretation of the bias in this simple setting, before moving to the more complicated multivariate setting in sections 3.2 and 3.3.

Consider the bias term of equation 2. This can be further simplified if we invert the $(X_1'X_1)$ and multiply by $X_1'X_2\beta_2$. Given that each of x_1 and x_2 are binary, we denote N_{x_1} and N_{x_2} as the quantity of observations for which x_1 and x_2 equal 1 respectively. This results in the following matrices for $X_1'X_1$ and $X_1'X_2$:

$$(X_1'X_1) = \begin{bmatrix} N & N_{x_1} \\ N_{x_1} & N_{x_1} \end{bmatrix} \qquad (X_1'X_2) = \begin{bmatrix} N_{x_2} \\ 0 \end{bmatrix},$$

and inverting $(X_1'X_1)$ gives:

$$(X_1'X_1)^{-1} = \begin{bmatrix} \frac{1}{N-N_{x_1}} & -\frac{1}{N-N_{x_1}} \\ -\frac{1}{N-N_{x_1}} & \frac{N}{N_{x_1}(N-N_{x_1})} \end{bmatrix}.$$

[2] This is laid out explicitly in Duflo et al. (2007) who states "If a researcher is cross-cutting interventions A and B, each of which has a comparison group, she obtains four groups: no interventions (pure control); A only; B only; and A and B together (full intervention)" It is important to note, that these four groups are mutually exclusive.

From equation 2, we can thus express the OVB formula for the included explanatory variable x_1 as[3] :

$$E[\widehat{\beta}_1^{ovb}|X] = \beta_1 - \beta_2 \frac{N_{x_2}}{N - N_{x_1}}. \tag{3}$$

This simple omitted variable bias also has a simple interpretation when cast in terms of treatment effect models. If x_2 (a treatment indicator) is omitted from the model, this group will be confounded with the true controls. Thus, the treatment effect on x_1 will be biased by any non-zero impact of x_2 on y (β_2), multiplied by the degree to which these x_2 units dilute the true control group: $N_{x_2}/(N - N_{x_1})$.

3.2 An Arbitrary Quantity of Omitted and Excluded Variables

While this simple OVB formula is intuitive, it is more interesting to be able to generalise this representation to a case with multiple omitted and included variables, which are much less frequently amenable to a clear interpretation using the original OVB formula, and simple economic logic. To do so, we extend to a case with an arbitrary quantity of included and excluded variables. Here we denote each of the k_1 variables in X_1 as $x_1^k \forall k = 1, \ldots, k_1$, (a constant is also included in X_1) and similarly, each of the k_2 variables in X_2 as $x_2^k \forall k = 1, \ldots, k_2$. Thus, $(X_1'X_1)$ and $(X_1'X_2)$ are given:

$$(X_1'X_1) = \begin{bmatrix} N & N_{x_1^1} & N_{x_1^2} & \cdots & N_{x_1^{k_1}} \\ N_{x_1^1} & N_{x_1^1} & 0 & \cdots & 0 \\ N_{x_1^2} & 0 & N_{x_1^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ N_{x_1^{k_1}} & 0 & 0 & \cdots & N_{x_1^{k_1}} \end{bmatrix} \qquad (X_1'X_2) = \begin{bmatrix} N_{x_2^1} & N_{x_2^2} & \cdots & N_{x_2^{k_2}} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Here the matrix $(X_1'X_1)$ consists of non-zero entries its main diagonal, first column and first row, and zeros elsewhere. The general class of matrices of this form (arrowhead matrices) has a simple inverse formula (Najafi et al., 2014), and based on this formula (refer to the Supplemental Appendix of the

[3]Note that given that this is a univariate regression model, this can also be derived using simply the covariance and variance, rather than matrices. In this case, we start with the simple bivariate version of the OVB. As each variable is binary, the covariance and variance have the simple closed form solutions below, where $N_{x_1x_2}$ refers to the quantity of observations for which both $x_1 = 1$ and $x_2 = 1$ (which is 0). As in 3, this gives:

$$E\widehat{\beta}_1^{ovb}|X] = \beta_1 + \beta_2 \left(\frac{Cov(x_1, x_2)}{Var(x_1)} \right) = \beta_1 + \beta_2 \left(\frac{(N_{x_1x_2} \cdot N - N_{x_1}N_{x_2})/N^2}{N_{x_1}(N - N_{x_1})/N^2} \right) = \beta_1 + \beta_2 \left(\frac{N_{x_2}}{N - N_{x_1}} \right).$$

paper for full algebra), the inverse of $(X_1'X_1)$ is given as:

$$(X_1'X_1)^{-1} = \begin{bmatrix} \frac{1}{\lambda} & -\frac{1}{\lambda} & -\frac{1}{\lambda} & \cdots & -\frac{1}{\lambda} \\ -\frac{1}{\lambda} & \frac{1}{N_{x_1^1}} + \frac{1}{\lambda} & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\ -\frac{1}{\lambda} & \frac{1}{\lambda} & \frac{1}{N_{x_1^2}} + \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{\lambda} & \frac{1}{\lambda} & \frac{1}{\lambda} & \cdots & \frac{1}{N_{x_1^{k_1}}} + \frac{1}{\lambda} \end{bmatrix}, \tag{4}$$

where $\lambda = N - N_{x_1^1} - N_{x_1^2} - \cdots - N_{x_1^{k_1}}$.

This leads to the multivariate generalisation of 3. In this case, each of the coefficients on the k_1 included variables has the expected value:

$$E[\widehat{\beta}_1^{k,ovb}|X] = \beta_1^k - \beta_2^1\left(\frac{N_{X_2^1}}{\lambda}\right) - \beta_2^2\left(\frac{N_{X_2^2}}{\lambda}\right) - \cdots - \beta_2^{k_2}\left(\frac{N_{X_2^{k_2}}}{\lambda}\right). \tag{5}$$

We note two things about this omitted variable bias. Firstly, it is identical for each of β_1^k terms. Secondly, as in the univariate case, this has an intuitive explanation when cast in terms of treatment effects. The treatment effect on each included variable will be biased by any non-zero impact of each excluded treatment group (the β_2^k terms), multiplied by the degree that each of these omitted treatment indicators biases the formation of the control group ($N_{x_2^k}/(N - N_{x_1^1} - \cdots - N_{x_1^{k_1}})$). In certain circumstances this formula will lead to clear upper or lower bounds on treatment effects. If each omitted treatment effect has the same sign as included treatment effects (this may be the case, for example, in spillovers of environmental shocks), estimated effects will be universally attenuated. Alternatively, if excluded treatments have opposite effects to included treatments, estimated effects will be consistently overstated in magnitude.

3.3 Extension to Alternative Treatment Models

Equation 5 provides an intuitive solution to the omitted variable bias in treatment effect models with multiple treatments, however only holds under the somewhat restrictive assumptions that each treatment (each variable) is mutually exclusive. Here we show that the omitted variable bias derived above can be logically extended to certain important extensions. To do this, we focus on the case of difference-in-differences style models. Consider the correctly specified model:

$$y_{i,t} = \beta_0 + \beta_t t + \beta_1 X_{1i,1} + \beta_2 X_{1i,t} + \gamma_1 X_{2i,1} + \gamma_2 X_{2i,t} + \varepsilon_{i,t} \tag{6}$$

Here we add a temporal component $t \in \{0, 1\}$, where observations are observed in two periods. Treatment takes a value of 1 for a subset of observations in time period $t = 1$, and 0 for all observations in time period $t = 0$. Treatment effects are thus estimated using time-varying indicators $X_{1i,t}$ and $X_{2i,t}$, and baseline differences in treated and untreated individuals are captured by the fixed effect $X_{1i,1}$ and $X_{2i,1}$. Any generalised temporal impacts are captured by the time fixed effect β_t. This multi-group difference-in-differences model is similar to that laid out in Imbens and Wooldridge (2009).

Suppose we estimate this model, omitting the vectors $X_{2i,t}$ and corresponding fixed effect $X_{2i,1}$. We denote this vector $X_2 = [X_{2i,1} \quad X_{2i,t}]$, and denote the vector $X_1 = [1 \quad X_{1i,1} \quad X_{1i,t} \quad t]$. We define k_1 as the quantity of treatment variables $X_{1i,t}$ in X_1, and k_2 the quantity of treatment variables in X_2. Here, the traditional OVB formula is given as:

$$E[\hat{\beta}^{ovb}|X] = \beta + (X_1'X_1)^{-1}X_1'X_2\gamma. \tag{7}$$

In what remains of this section I document the same convenient OVB formula as in sections 3.1-3.2 for this difference-in-differences model. To do so, first I consider the case of a single included and omitted treatment variable (and corresponding fixed effects), before extending to the case with an arbitrary quantity of included and omitted variables.

Given that X_1 is a Boolean matrix, and various columns are mutually exclusive, we can write $X_1'X_1$ and $X_1'X_2$ as:

$$(X_1'X_1) = \begin{bmatrix} N & N_{x_1} & N_{x_{1t}} & N_t \\ N_{x_1} & N_{x_1} & 0 & 0 \\ N_{x_{1t}} & 0 & N_{x_{1t}} & N_{x_{1t}} \\ N_t & 0 & N_{x_{1t}} & N_t \end{bmatrix} \qquad X_1'X_2 = X_1'[x_{2i,1} \quad X_{2i,t}] = \begin{bmatrix} N_{x_2} & N_{x_{2t}} \\ 0 & 0 \\ 0 & 0 \\ 0 & N_{x_{2t}} \end{bmatrix}.$$

As before, N refers to the number of observations, N_{x1} to those for which $X_{1,i1} = 1$, N_{x1t} to the quantity for which $X_{1,it} = 1$, and similarly for N_{x2} and N_{x2t}. Additionally N_t is equal to the number of observations for which $t = 1$.

Although the $(X_1'X_1)$ matrix is now more complex than in the static treatment model, once again we can show that it has a reasonably simple closed form solution for the inverse. As it is a 2×2 block matrix, it can be inverted (full details are available in the Supplemental Appendix of this paper and Lu and Shiou (2002)) resulting in:

$$
(X_1'X_1)^{-1} = \begin{bmatrix} \frac{1}{N-N_t-N_{x_1}} & -\frac{1}{N-N_t-N_{x_1}} & 0 & 0 \\ -\frac{1}{N-N_t-N_{x_1}} & \frac{N-N_t}{N_{x_1}(N-N_t-N_{x_1})} & -\frac{1}{N-N_t-N_{x_1}} & \frac{1}{N-N_t-N_{x_1}} \\ 0 & -\frac{1}{(N-N_{x_1}-N_t)} & \frac{N_t}{N_{x_{1t}}(N_t-N_{x_{1t}})} & \frac{-1}{N_t-N_{x_{1t}}} \\ 0 & \frac{1}{(N-N_{x_1}-N_t)} & \frac{-1}{N_t-N_{x_{1t}}} & \frac{N-N_{x_1}-N_{x_{1t}}}{(N-N_{x_1}-N_t)(N_t-N_{x_{1t}})} \end{bmatrix}. \tag{8}
$$

Putting this together following equation 7 results in the following omitted variable formula for the time-varying element of a difference-in-differences model:

$$
E[\widehat{\beta}_2^{ovb}|X] = \beta_2 - \gamma_2\left(\frac{N_{x_{2t}}}{N_t - N_{x_{2T}}}\right). \tag{9}
$$

In this case, we thus have that the OVB follows virtually the same logic as in equation 3, however now conditional on treatment occurring in the second period. The omission of a relevant treatment indicator in difference-in-differences models thus biases included treatment effects by any effect that this treatment indicator has on the outcome of interest, multiplied by the proportion of the "control group" that are actually treated, and hence should have been included in the regression.

Finally, note that this OVB formula can be resolved in this way even in the extreme case of multiple included and multiple omitted treatment indicators. To see this, we return to equation 6, where both $k_1 > 1$ and $k_2 > 1$. Following the notation above, we can thus write the $X_1'X_1$ (at left) and $X_1'X_2$ (at right) matrices as:

$$
\begin{bmatrix}
N & N_{x_1^1} & N_{x_1^2} & \dots & N_{x_1^{k_1}} & N_{x_{1t}^1} & N_{x_{1t}^2} & \dots & N_{x_{1t}^{k_1}} & N_t \\
N_{x_1^1} & N_{x_1^1} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
N_{x_1^2} & 0 & N_{x_1^2} & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
N_{x_1^{k_1}} & 0 & 0 & \dots & N_{x_1^{k_1}} & 0 & 0 & \dots & 0 & 0 \\
N_{x_{1t}^1} & 0 & 0 & \dots & 0 & N_{x_{1t}^1} & 0 & \dots & 0 & N_{x_{1t}^1} \\
N_{x_{1t}^2} & 0 & 0 & \dots & 0 & 0 & N_{x_{1t}^2} & \dots & 0 & N_{x_{1t}^2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
N_{x_{1t}^{k_1}} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & N_{x_{1t}^{k_1}} & N_{x_{1t}^{k_1}} \\
N_t & 0 & 0 & \dots & 0 & N_{x_{1t}^1} & N_{x_{1t}^2} & \dots & N_{x_{1t}^{k_1}} & N_t
\end{bmatrix}
\qquad
\begin{bmatrix}
X_2^1 & \dots & X_2^{k_2} & X_{2t}^1 & \dots & X_{2t}^{k_2} \\
0 & \dots & 0 & 0 & \dots & 0 \\
0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 \\
0 & \dots & 0 & 0 & \dots & 0 \\
0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 \\
0 & \dots & 0 & X_{2t}^1 & \dots & X_{2t}^{k_2}
\end{bmatrix}.
$$

Once again, $X_1'X_1$ is a 2×2 block matrix, with each block of dimension $(1+k_1)\times(1+k_1)$. The blocks on the principal diagonal are each arrowhead matrices, and have a known inverse, and the blocks on the off-diagonal have values in the first row and column respectively, with the remainder of entries equal to zero. Thus, each block of the $X_1'X_1$ matrix is invertible, as is the underlying matrix. Full algebra is available in Supplemental Information to this paper. The inverse of $X_1'X_1$ is thus written:

$$(X_1'X_1)^{-1} = \begin{bmatrix} \frac{1}{\theta} & \frac{1}{\theta} & \frac{1}{\theta} & \cdots & -\frac{1}{\theta} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\theta} & \frac{\theta+N_{x_1^1}}{\theta N_{x_1^1}} & \frac{1}{\theta} & \cdots & -\frac{1}{\theta} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\theta} & \frac{1}{\theta} & \frac{\theta+N_{x_1^2}}{\theta N_{x_1^2}} & \cdots & -\frac{1}{\theta} & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{\theta} & -\frac{1}{\theta} & -\frac{1}{\theta} & \cdots & \frac{\theta+N_{x_1^{k_1}}}{\theta N_{x_1^{k_1}}} & -\frac{1}{\theta} & -\frac{1}{\theta} & \cdots & -\frac{1}{\theta} & \frac{1}{\theta} \\ 0 & 0 & 0 & \cdots & -\frac{1}{\theta} & \frac{\theta_t+N_{x_{1t}^1}}{\theta_t N_{x_{1t}^1}} & \frac{1}{\theta_t} & \cdots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\ 0 & 0 & 0 & \cdots & -\frac{1}{\theta} & \frac{1}{\theta_t} & \frac{\theta_t+N_{x_{1t}^2}}{\theta_t N_{x_{1t}^2}} & \cdots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{\theta} & \frac{1}{\theta_t} & \frac{1}{\theta_t} & \cdots & \frac{\theta_t+N_{x_{1t}^{k_1}}}{\theta_t N_{x_{1t}^{k_1}}} & -\frac{1}{\theta_t} \\ 0 & 0 & 0 & \cdots & \frac{1}{\theta} & -\frac{1}{\theta_t} & -\frac{1}{\theta_t} & \cdots & -\frac{1}{\theta_t} & \frac{\theta_t+\theta}{\theta_t\theta} \end{bmatrix} \tag{10}$$

where $\theta = N - N_t - N_{x_1^1} - N_{x_1^2} - \cdots - N_{x_1^{k_1}}$ and $\theta_t = N_t - N_{x_{1t}^1} - N_{x_{1t}^2} - \cdots - N_{x_{1t}^{k_1}}$. Note that the inverse in equation 8 is just a special case of the above, where $k_1 = 1$. Now, to determine the omitted variable bias *on each* included variable where potentially multiple treatment variables are included and excluded in the difference-in-differences model, we return to equation 7. From this, and the preceding matrices, we have that the OVB formula on each treatment effect in time-varying $X_{1i,t}$ is:

$$E[\widehat{\beta}_2^{k,ovb}|X] \;=\; \beta_2^k - \gamma_2^1\frac{N_{x_{2t}^1}}{\theta_t} - \gamma_2^2\frac{N_{x_{2t}^2}}{\theta_t} - \gamma_2^{k_2}\frac{N_{x_{2t}^{k_2}}}{\theta_t} \quad \forall k \in 1,\ldots,k_1. \tag{11}$$

This bias term is once again intuitive in the treatment effects framework. Here each binary treatment indicator which was omitted from the estimation is incorrectly included in the control group, and hence produces a bias in the estimated treatment effects of included variables. This bias consists of the true treatment effect of non-included treatment variables (γ_2^k) scaled by the degree to which this treatment group contaminates the naive control group ($N_{x_{2t}^k}/(N_t - N_{x_{1t}^1} - N_{x_{1t}^2} - \cdots - N_{x_{1t}^{k_1}})$). If a researcher has prior information to suggest that non-included treatment units are related in some way to included treatment units, this may allow an even finer consideration of this bias. As discussed previously, if the impact of treatment is of the same direction on included and non-included treatment units, the derivation in 11 proves an attenuation bias, while if non-included treatment units are thought to be of the *opposite* direction as the included treatment units, estimates from the incorrectly specified model will consistently overstate the true treatment impacts on included variables.

4 Discussion and Conclusion

The Omitted Variable Bias is frequently encountered in economics. While it is the base of a range of useful derivations, in the case where multiple omitted variables are considered in regressions, it is often presented as an *ex-post* test of model stability, rather than as providing a simple *ex-ante* formula for determining parameter bounds. An *ex-post* derivation of this type is provided by Gelbach (2016).

In this paper I provide a simple and intuitive formula for the OVB in treatment effects models, where a variable of interest is regressed on multiple binary treatment variables (and subject to multiple other omitted treatment variables). This OVB is likely to be particularly convenient in situations where multiple treatments or multiple levels of treatment exist, but assignation to treatment is imperfectly observed. Such cases are particularly common in natural experiments where treatment is not under the precise control of a human experimenter. As these natural experiments are often analysed using difference-in-differences models, I document that this simple OVB formula has a logical extension to these models, without complicating the intuition (or the algebra) behind the formula.

I show that the OVB formula consists of two elements: the degree to which omitted treatment indicators are incorrectly included in the control group of regression models, and the impact that these omitted treatment indicators have on outcomes of interest. This is, by definition, the counterpart to the textbook OVB formula, however potentially leads to much simpler formation of hypotheses regarding the degree or magnitude of this omitted bias.

References

D. Almond, L. Edlund, and M. Palme. Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. *The Quarterly Journal of Economics*, 124(4): 1729–1772, 2009.

A. V. Banerjee, S. Cole, E. Duflo, and L. Linden. Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics*, 122(3):1235–1264, 2007.

G. Baron, E. Perrodeau, I. Boutron, and P. Ravaud. Reporting of analyses from randomized controlled trials with multiple arms: a systematic review. *BMC Medicine*, 11(84):1–7, 2013.

K. A. Clarke. The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management and Peace Science*, 22(4):341–352, May 2005.

E. Duflo, R. Glennerster, and M. Kremer. Chapter 61 Using Randomization in Development Economics Research: A Toolkit. In T. P. Schultz and J. A. Strauss, editors, *Handbook of Development Economics*, volume 4, pages 3895 – 3962. Elsevier, 2007.

J. B. Gelbach. When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics*, 34(2):509–543, 2016.

W. H. Greene. *Econometric Analysis*. Prentice Hill, New Jersey, 5 edition, 2002.

G. W. Imbens and J. M. Wooldridge. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1):5–86, 2009.

M. Kremer. Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. *American Economic Review*, 93(2):102–106, May 2003.

T.-T. Lu and S.-H. Shiou. Inverses of 2×2 Block Matrices. *Computers & Mathematics with Applications*, 43(1–2):119–129, 2002.

E. Miguel and M. Kremer. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1):159–217, 01 2004.

H. S. Najafi, S. A. Edalatpanah, and G. A. Gravvanis. An efficient method for computing the inverse of arrowhead matrices. *Applied Mathematics Letters*, 33(1):1–5, 2014.

H. Theil. Specification Errors and the Estimation of Economic Relationships. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 25(1/3):41–51, 1957.

SUPPLEMENTAL APPENDIX
A Convenient Omitted Variable Bias Formula for Treatment Effect Models

Damian Clarke

A Inverting the Arrowhead Matrix in the Multivariate Treatment Effects Model

In equation 4 of the paper, the inverse of the $X_1'X_1$ matrix is found following the known inverse formula for arrowhead matrices. This formula assumes a $J \times J$ arrowhead matrix of the form:

$$A = \begin{bmatrix} \alpha & z' \\ z & D \end{bmatrix}$$

where α is a scalar, z a vector with $J-1$ elements, and D a diagonal $J-1 \times J-1$ matrix with zeros on the off-diagonal. If each of the diagonal terms is non-zero (an assumption which is met by construction provided that each treatment variable has at least treated observation), the inverse is equal to

$$A^{-1} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & D^{-1} \end{bmatrix} + \rho u u'$$

where $u = [-1 \quad D^{-1}z]'$ and $\rho = \frac{1}{\alpha - z'D^{-1}z}$.

In this case, the $X_1'X_1$ matrix is equal to:

$$(X_1'X_1) = \begin{bmatrix} N & N_{x_1^1} & N_{x_1^2} & \cdots & N_{x_1^{k_1}} \\ N_{x_1^1} & N_{x_1^1} & 0 & \cdots & 0 \\ N_{x_1^2} & 0 & N_{x_1^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ N_{x_1^{k_1}} & 0 & 0 & \cdots & N_{x_1^{k_1}} \end{bmatrix}$$

implying that $\alpha = N$, $z = [N_{x_1^1} \quad N_{x_1^2} \quad \cdots \quad N_{x_1^{k_1}}]'$ and D is the matrix contained in from row 2 to

J and column 2 to J of $(X_1'X_1)$. Putting this together gives that:

$$(X_1'X_1)^{-1} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1/N_{x_1^1} & 0 & \cdots & 0 \\ 0 & 0 & 1/N_{x_1^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1/N_{x_1^{k_1}} \end{bmatrix} + \frac{1}{\lambda} \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

where $\lambda = N - N_{x_1^1} - N_{x_1^2} - \cdots - N_{x_1^{k_1}}$, and resolving gives:

$$(X_1'X_1)^{-1} = \begin{bmatrix} \frac{1}{\lambda} & -\frac{1}{\lambda} & -\frac{1}{\lambda} & \cdots & -\frac{1}{\lambda} \\ -\frac{1}{\lambda} & \frac{1}{N_{x_1^1}} + \frac{1}{\lambda} & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\ -\frac{1}{\lambda} & \frac{1}{\lambda} & \frac{1}{N_{x_1^2}} + \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{\lambda} & \frac{1}{\lambda} & \frac{1}{\lambda} & \cdots & \frac{1}{N_{x_1^{k_1}}} + \frac{1}{\lambda} \end{bmatrix},$$

as documented in equation 4.

B Inverting the 2×2 Block Matrix for OVB in Simple Difference-in-Differences

In order to calculate the omitted variable bias in a difference-in-differences model with one included and one excluded treatment indicator (plus corresponding fixed effects), we must invert $X_1'X_1$, as displayed in equation 8. This matrix is a symmetric block matrix, and so can be re-written as:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{bmatrix}$$

where each of A_{11}, A_{12} and A_{22} are the 2×2 matrices in each corner of $(X_1'X_1)$. The formula for the inverse of a 2×2 block matrix, and the algebra corresponding to $X_1'X_1$ implies that each element of the inverse has the formula given below (see for example Lu and Shiou (2002)):

$$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{12}')^{-1}$$

$$= \frac{1}{N_{x_1}(N - N_t - N_{x_1})} \begin{bmatrix} N_{x_1} & -N_{x_1} \\ -N_{x_1} & N - N_t \end{bmatrix} \tag{12}$$

$$\begin{aligned} B_{22} &= (A_{22} - A'_{12}A_{11}^{-1}A_{12})^{-1} \\ &= \frac{1}{(N_t - N_{x_{1t}})} \begin{bmatrix} \frac{N_t}{N_{x_{1t}}} & -1 \\ -1 & \frac{N - N_{x_1} - N_{x_{1t}}}{(N - N_{x_1} - N_t)} \end{bmatrix} \end{aligned} \tag{13}$$

$$\begin{aligned} B_{12} &= -A_{22}^{-1}A'_{12}(A_{11} - A_{12}A_{22}^{-1}A'_{12})^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ -\frac{1}{N - N_t - N_{x_1}} & \frac{1}{N - N_t - N_{x_1}} \end{bmatrix} \end{aligned} \tag{14}$$

$$\begin{aligned} B'_{12} &= -A_{11}^{-1}A_{12}(A_{22} - A'_{12}A_{11}^{-1}A_{12})^{-1} \\ &= \begin{bmatrix} 0 & -\frac{1}{(N - N_{x_1} - N_t)} \\ 0 & \frac{1}{(N - N_{x_1} - N_t)} \end{bmatrix}. \end{aligned} \tag{15}$$

The first line of each expression above is from the inverse formula for 2×2 block matrices, while the second line for each sub-matrix is resolved by linear algebra.Putting the four elements of B together gives:

$$A^{-1} = (X'_1X_1)^{-1} = \begin{bmatrix} \frac{1}{N - N_t - N_{x_1}} & -\frac{1}{N - N_t - N_{x_1}} & 0 & 0 \\ -\frac{1}{N - N_t - N_{x_1}} & \frac{N - N_t}{N_{x_1}(N - N_t - N_{x_1})} & -\frac{1}{N - N_t - N_{x_1}} & \frac{1}{N - N_t - N_{x_1}} \\ 0 & -\frac{1}{(N - N_{x_1} - N_t)} & \frac{N_t}{N_{x_{1t}}(N_t - N_{x_{1t}})} & \frac{-1}{N_t - N_{x_{1t}}} \\ 0 & \frac{1}{(N - N_{x_1} - N_t)} & \frac{-1}{N_t - N_{x_{1t}}} & \frac{N - N_{x_1} - N_{x_{1t}}}{(N - N_{x_1} - N_t)(N_t - N_{x_{1t}})} \end{bmatrix}.$$

as indicated in equation 8 in the paper.

C Inverting the X'_1X_1 Matrix in a Difference-in-Differences Model with Multiple Included and Excluded Treatment Indicators

In a case with k_1 included treatment indicators, determining the inverse of the X'_1X_1 matrix requires inverting a $(2 + 2k_1) \times (2 + 2k_1)$ matrix, based on the k_1 treatment indicators, and identical number of fixed effects, an intercept term, and a time dummy. Given that this is once again a 2×2 block matrix, we can use the identical formula as in section B for the inverse of each block of $X_1' X_1$. However, now each block is of dimension $(k_1+1) \times (k_1+1)$. Fortunately, the blocks on the principal diagonal are arrowhead matrices, and the blocks on the off-diagonal have elements only in row 1 or column 1, and zeros elsewhere. Thus in each case, the required matrices are easily invertible using the arrowhead matrix formula (see Najafi et al. (2014) for discussion), and in the case of B_{22}, the final inverse is found using the bordering method for symmetric matrices.

Each element of the inverse is documented below:

$$
\begin{aligned}
B_{11} &= (A_{11} - A_{12}A_{22}^{-1}A_{12}')^{-1} = \begin{bmatrix}
\frac{1}{\theta} & \frac{1}{\theta} & \frac{1}{\theta} & \cdots & -\frac{1}{\theta} \\
\frac{1}{\theta} & \frac{\theta + N_{x_1^1}}{\theta_0 N_{x_1^1}} & \frac{1}{\theta} & \cdots & -\frac{1}{\theta} \\
\frac{1}{\theta} & \frac{1}{\theta} & \frac{\theta + N_{x_1^2}}{\theta_0 N_{x_1^2}} & \cdots & -\frac{1}{\theta} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\frac{1}{\theta} & -\frac{1}{\theta} & -\frac{1}{\theta} & \cdots & \frac{\theta + N_{x_1^{k_1}}}{\theta N_{x_1^{k_1}}}
\end{bmatrix} \qquad (16) \\
\text{where } \theta &= N - N_t - N_{x_1^1} - N_{x_1^1} - \cdots - N_{x_1^{k_1}},
\end{aligned}
$$

and

$$
\begin{aligned}
B_{22} &= (A_{22} - A_{12}'A_{11}^{-1}A_{12})^{-1} = \begin{bmatrix}
\frac{\theta_t + N_{x_{1t}^1}}{\theta_t N_{x_{1t}^1}} & \frac{1}{\theta_t} & \cdots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\
\frac{1}{\theta_t} & \frac{\theta_t + N_{x_{1t}^2}}{\theta_t N_{x_{1t}^2}} & \cdots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{1}{\theta_t} & \frac{1}{\theta_t} & \cdots & \frac{\theta_t + N_{x_{1t}^{k_1}}}{\theta_t N_{x_{1t}^{k_1}}} & -\frac{1}{\theta_t} \\
-\frac{1}{\theta_t} & -\frac{1}{\theta_t} & \cdots & -\frac{1}{\theta_t} & \frac{\theta + \theta_t}{\theta_t \theta}
\end{bmatrix} \qquad (17) \\
\text{where } \theta_t &= N_t - N_{x_{1t}^1} - N_{x_{1t}^2} - \cdots - N_{x_{1t}^{k_1}}
\end{aligned}
$$

for the diagonal blocks, and

$$
B_{12} = -A_{22}^{-1}A_{12}'(A_{11} - A_{12}A_{22}^{-1}A_{12}')^{-1} = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 \\
-\frac{1}{\theta} & -\frac{1}{\theta} & -\frac{1}{\theta} & \cdots & \frac{1}{\theta}
\end{bmatrix} \qquad (18)
$$

and

$$B'_{12} \;=\; -A_{11}^{-1}A_{12}(A_{22}-A'_{12}A_{11}^{-1}A_{12})^{-1} = \begin{bmatrix} 0 & 0 & \dots & 0 & -\frac{1}{\theta} \\ 0 & 0 & \dots & 0 & -\frac{1}{\theta} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -\frac{1}{\theta} \\ 0 & 0 & \dots & 0 & \frac{1}{\theta} \end{bmatrix}. \tag{19}$$

for the off-diagonal terms.

Combining 16-19 gives the following for $(X'_1X_1)^{-1}$:

$$(X'_1X_1)^{-1} = \begin{bmatrix} \frac{1}{\theta} & \frac{1}{\theta} & \frac{1}{\theta} & \dots & -\frac{1}{\theta} & 0 & 0 & \dots & 0 & 0 \\ \frac{1}{\theta} & \frac{\theta+N_{x_1^1}}{\theta N_{x_1^1}} & \frac{1}{\theta} & \dots & -\frac{1}{\theta} & 0 & 0 & \dots & 0 & 0 \\ \frac{1}{\theta} & \frac{1}{\theta} & \frac{\theta+N_{x_1^2}}{\theta N_{x_1^2}} & \dots & -\frac{1}{\theta} & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{\theta} & -\frac{1}{\theta} & -\frac{1}{\theta} & \dots & \frac{\theta+N_{x_1^{k_1}}}{\theta N_{x_1^{k_1}}} & -\frac{1}{\theta} & -\frac{1}{\theta} & \dots & -\frac{1}{\theta} & \frac{1}{\theta} \\ 0 & 0 & 0 & \dots & -\frac{1}{\theta} & \frac{\theta_t+N_{x_{1t}^1}}{\theta_t N_{x_{1t}^1}} & \frac{1}{\theta_t} & \dots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\ 0 & 0 & 0 & \dots & -\frac{1}{\theta} & \frac{1}{\theta_t} & \frac{\theta_t+N_{x_{1t}^2}}{\theta_t N_{x_{1t}^2}} & \dots & \frac{1}{\theta_t} & -\frac{1}{\theta_t} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\frac{1}{\theta} & \frac{1}{\theta_t} & \frac{1}{\theta_t} & \dots & \frac{\theta_t+N_{x_{1t}^{k_1}}}{\theta_t N_{x_{1t}^{k_1}}} & -\frac{1}{\theta_t} \\ 0 & 0 & 0 & \dots & \frac{1}{\theta} & -\frac{1}{\theta_t} & -\frac{1}{\theta_t} & \dots & -\frac{1}{\theta_t} & \frac{\theta_t+\theta}{\theta_t\theta} \end{bmatrix}$$

as laid out in equation 10 of the paper.