Exogeneity and Robustness

Bryant Chen and Judea Pearl
University of California, Los Angeles
Computer Science Department
Los Angeles, CA, 90095-1596, USA
(310) 825-3243
February 22, 2015

Abstract

A common practice for detecting misspecification is to perform a “robustness test”, where the researcher examines how a regression coefficient of interest behaves when variables are added to the regression. Robustness of the regression coefficient is taken as evidence of structural validity. However, there are numerous pitfalls that can befall a researcher when performing such tests. For example, we demonstrate that certain regressors, when added to the regression, will induce a shift in the coefficient of interest even when structurally valid. Such robustness tests would produce false alarm, suggesting that the model is misspecified when it is not. For a robustness test to be informative, the variables added to the regression must be carefully chosen based on the model structure. We provide a simple criterion that allows researchers to quickly determine which variables, when added to the regression, constitute informative robustness tests. We also explore the extent to which robustness tests are able to detect bias, demonstrating that robustness tests enable detection of bias due not only to omitted observable variables but omitted unobservable variables as well. Finally, we empirically verify many of the results derived using Monte Carlo simulations.

1 Introduction

Suppose that economic theory dictates the structural equation

\[ Y = \beta_0 D + \alpha_0 Z + U, \]

where \( D \) is an observable cause of interest, \( Y \) is an outcome of interest, \( Z \) is an observable driver of \( Y \), and \( U \) represents unobservable drivers of \( Y \), often called a “disturbance” or “error” term\(^1\). To estimate the value of \( \beta_0 \), we may regress \( Y \) on \( D \), and a set of variables, \( W = \{W_1, W_2, ..., W_k\} \), where \( Z \subseteq W \). This gives the regression equation

\[ Y = r_0 D + r_1 W_1 + r_2 W_2 + ... + r_k W_k + \epsilon, \]

\(^1\)\( D \) and \( Z \) may also be vectors of variables without changing any of the results in this paper but for convenience we assume that they are singletons.
where $\epsilon$ is the regression residual and $r_0$ is the partial regression coefficient of $D$ when $Y$ is regressed on $D$ and $W$, that is, $r_0 = r_{YD,W}$.

It is well known that $\beta_0$ is identified and equal to $r_{YD,W}$ if $W$ satisfies the conditional independence $D \perp U | W$, often called the conditional exogeneity assumption. If this assumption fails then the regression coefficient $r_{YD,W}$ will generally not equal $\beta_0$. Unfortunately, since $U$ is unobservable, it is impossible to test whether this assumption holds, and the task of defending the validity of $D \perp U | W$ is delegated to human judgment, which is vulnerable to two sources of error. The first, of course, is model misspecification: although $D \perp U | W$ can in principle be verified from the hypothesized structure, one can never be sure whether these theoretical assumptions are valid. Second, to judge whether $D \perp U | W$ holds in a given specified model can be formidable when multiple equations and multiple $U$ factors are present, some correlated with the observable variables and some with other $U$ factors.

While the conditional exogeneity assumption is not directly testable, one can identify various implications of the model structure that can be used to test it against data. If these implications are found compatible with the data then the model gains credibility. Consequently, if the model implies conditional exogeneity then it also gains credibility. Testing the conditional exogeneity assumption along these lines requires two steps. First, to identify and test all testable implications of the model, and second, to verify that conditional exogeneity holds in the model. The first step can be rather involved and is rarely performed in practice. Instead, practitioners resort to shortcuts—testing only a subset of implications deemed relevant to the conditional exogeneity assumption.

A common exercise in empirical studies, which utilizes such shortcuts, is to check the “robustness” of certain regression coefficients when the regression specification is modified by including or excluding “control” covariates. Movement by the regression coefficients of $D$ is then taken as evidence of omitted variable bias or misspecification. In a recent survey of non-experimental empirical work, Lu and White (2014) found that of the 76 papers involving data analysis published in The American Economic Review during 2009, 23 perform a robustness check along the lines just described. Similarly, Oster (2013) found that 75% of 2012 papers published in The American Economic Review, Journal of Political Economy, and Quarterly Journal of Economics explored the sensitivity of results to varying control sets.

The intuition behind this procedure is rooted on the following heuristic: if bias is caused by some set $W$ of confounders, then controlling for $W$ (by adding $W$ to the regression equation) should eliminate that bias. Any further control would then be unnecessary, and should leave $r_0$ unaltered. The invariance of $r_0$ to additional regressors is taken as evidence that all confounders have been accounted for, and none remain outside the set $W$. Conversely, if $r_0$ shifts with the addition of regressors beyond $W$, it is taken as evidence that $W$ was not sufficient to cover all confounders and, consequently, $r_0$ is likely to be biased and the conditional exogeneity $D \perp U | W$ is likely violated.

\[2\] The methods for testing conditional exogeneity proposed by White and Chalak (2010) can be understood in this manner. However, these methods require specific structural assumptions beyond conditional exogeneity and are, therefore, not applicable to all models.

\[3\] The practice of assessing missing variable bias by observing the sensitivity of an estimator to additional controls is not unique to economics. A recent survey of articles in major epidemiologic journals by Walter and Tiemeier (2009) found that 15% of papers used a “change-in-estimate” criterion to select covariates.
This heuristic assumes that bias reduction is monotonic with the number of regressors, that is, that adding regressors cannot create bias where none exists. We will show that this assumption is false. Certain regressors, if added to $W$ will necessarily change $r_0$, even when $W$ is sufficient to satisfy the exogeneity condition $D \perp U | W$. We call such regressors shift-producing. Robustness tests involving shift-producing covariates are non-informative and produce false alarm when exogeneity holds. We will show how shift-producing regressors can be identified from the model's structure.

For informative tests, the connection between robustness and exogeneity is as follows. If conditional exogeneity holds before and after the addition of regressors then obviously $r_0$ will be invariant. Conversely, if $r_0$ shifts then conditional exogeneity must be violated before or after the addition. As a result, if the hypothesized model structure dictates that exogeneity holds in both cases then the added regressor may be informative since failure of the corresponding robustness test indicates model misspecification. However, as mentioned previously, the task of determining whether exogeneity holds, even in a well-specified model, can be formidable. As formal and transparent representations of the model structure, causal graphs provide researchers with the means to determine, by inspection, whether a given set of variables satisfies the conditional exogeneity assumption. Utilizing a graphical condition, called single-door criterion, we are able to quickly identify sets, $W_1$ and $W_2$, for which the model implies $r_{YD,W_1} = r_{YD,W_1W_2} = \beta_0$. We will show that comparing the corresponding regressions, $r_{YD,W_1}$ and $r_{YD,W_1W_2}$, constitutes an informative robustness test in that failure both implies model misspecification and is possible when the conditional exogeneity assumption is violated. As a result, researchers can focus on the economic plausibility of the model structure and substitute all other judgments with sound and reliable mechanical procedures when identifying $\beta_0$ and finding informative covariates for robustness testing.

Graphs have been utilized by economists to communicate causal structure and facilitate economic problems since the 1930s (Tinbergen, 1939). Orcutt (1952), for instance, used graphs to represent possible causal structures for a given set of variables and gave examples illustrating that some graphs are incompatible with certain conditional independences among the variables (e.g. zero partial correlation between $Z$ and $X$ given $Y$ is incompatible with the chain $Y \rightarrow X \rightarrow Z$), thus allowing researchers to reject certain causal structures. Tinbergen (1939), Wold (1954), and other practitioners of process analysis also employed graphs to convey the causal relationship between variables (Hoover, 2004). More recently, graphs have been utilized by White and Chalak (2009), White and Lu (2011), and Hoover and Phiromswad (2013) to facilitate problems of identification, optimization, identifying instrumental variables from data, and more. We will use graphical models primarily to detect conditional independences and verify identifying assumptions.

Our paper is structured in the following way: We will begin Section 2 by introducing graphical representations of structural models with special interest on graph separation, a notion that will play a pivotal role in the results that follow. In Section 3, we introduce two graphical criteria that will enable us to identify structural parameters. The first, called single-door criterion, provides a necessary and sufficient condition for the identification of $\beta_0$ using regression. The second, called back-door criterion, provides a necessary and sufficient condition for the identification of the total effect of $D$ on $Y$ using regression. These tools will allow us to characterize shift-producing
Figure 1: (a) Model with latent variables \((Q_1 \text{ and } Q_2)\) shown explicitly (b) Same model with latent variables summarized

covariates and informative robustness tests in Section 4. We also explore the extent to which robustness tests are able to detect bias, demonstrating that robustness tests are to detect not only omitted observable variables but omitted unobservable variables as well. Finally, in Section 5, we empirically verify many of our findings in Section 4 using Monte Carlo simulations.

2 Preliminaries

2.1 Causal Graphs

We introduce causal graphs by way of example. Suppose we wish to estimate the effect of attending an elite college on future earnings. Clearly, simply regressing earnings on college rating will not give an unbiased estimate of the target effect. This is because elite colleges are highly selective so students attending them are likely to have qualifications for high-earning jobs prior to attending the school. This background knowledge can be expressed in the following model specification.

Model 1.

\[
\begin{align*}
Q_1 & = U_1 \\
C & = a \cdot Q_1 + U_2 \\
Q_2 & = c \cdot C + d \cdot Q_1 + U_3 \\
S & = b \cdot C + e \cdot Q_2 + U_4,
\end{align*}
\]

where \(Q_1\) represents the individual’s qualifications prior to college, \(Q_2\) represents qualifications after college, \(C\) contains attributes representing the quality of the college attended, and \(S\) the individual’s salary.

Figure 1a is a causal graph that represents this model specification. Each variable in the model has a corresponding node or vertex in the graph. Additionally, for each equation, arrows are drawn from the independent variables to the dependent variables. These arrows reflect the direction of causation. In some cases, we may label the arrow with its corresponding structural coefficient as in Figure 1a. Error terms are typically not displayed in the graph.
If $Q_1$ and $Q_2$ are unobservable or latent variables their influence on $S$ is generally attributed to $S$’s error term. By removing them, we obtain the following model specification:

**Model 2.**

\[
C = U_C \\
S = \beta C + U_S
\]

The background information specified by Model 1 imply that the error term of $S$, $U_S$, is not independent of $U_C$ and, as a result, not independent of $C$. As a result, exogeneity does not hold. Dependence between error terms is depicted in the causal graph as a bidirected arc between the variables whose error terms are correlated as in Figure 1b.

It is clear that $\beta$ is not identified in Model 2. However, if we include the strength of an individual’s college application, $A$, as shown in Figure 2a, we obtain the following model:

**Model 3.**

\[
Q_1 = U_1 \\
A = a \cdot Q_1 + U_2 \\
C = b \cdot A + U_3 \\
Q_2 = e \cdot Q_1 + d \cdot C + U_4 \\
S = c \cdot C + f \cdot Q_2 + U_5.
\]

By removing the latent variables from the model specification we obtain:

**Model 4.**

\[
A = a \cdot Q_1 + U_A \\
C = b \cdot A + U_C \\
S = \beta \cdot C + U_S,
\]
Figure 3: Model illustrating the rules of d-separation

Now, $\beta$ is identified and can be estimated using the regression of $S$ on $C$ and $A$. We will show how this and more complicated identification results can be obtained using the causal graph in Section 3.

In summary, the causal graph is constructed from the model equations in the following way: Each variable in the model has a corresponding vertex or node in the graph. For each equation, arrows are drawn in the graph from the dependent variables to the independent variable. Finally, if the error terms of any two variables are correlated, then a bidirected edge is drawn between the two variables.

Before continuing, we review some basic graph terminology. If an arrow, called $(X,Y)$, exists from $X$ to $Y$ we say that $X$ is a parent of $Y$. If there exists a sequence of arrows all of which are directed from $X$ to $Y$ we say that $X$ is an ancestor of $Y$. If $X$ is an ancestor of $Y$ then $Y$ is a descendant of $X$. The set of nodes connected to $Y$ by a bidirected arc are called the siblings of $Y$. Lastly, a collider is a node where colliding arrowheads meet. Z in Figure 4a is a collider as are $C$, $D$, and $E$ in Figure 3.

A path between $X$ to $Y$ is a sequence of edges, connecting the two vertices. A path may go either along or against the direction of the arrows. A directed path from $X$ to $Y$ is a path begins with an arrow pointing to $X$ and ends with an arrow pointing to $Y$. For example, in Figure 3, $C \leftarrow B \rightarrow E$, $C \rightarrow D \rightarrow E$, $C \leftarrow B \rightarrow D \rightarrow E$, and $C \rightarrow D \leftarrow B \rightarrow E$ are all paths between $C$ and $E$. However, only $C \rightarrow D \rightarrow E$ is a directed path, and only $C \leftarrow D \rightarrow E$ and $C \leftarrow B \rightarrow D \rightarrow E$ are back-door paths. The significance of directed paths stems from the fact that they convey the flow of causality, while the significance of back-door paths stems from their association with confounding.

A graph is acyclic if it does not contain any cycles, a directed path that begins and ends with the same node. A graph is cyclic if it contains a cycle. A model in which the causal graph is acyclic is called recursive while models with cyclic graphs are called non-recursive.

2.2 D-Separation

D-separation allows researchers to identify conditional independences implied by the model’s structure from the causal graph. The idea of d-separation is to associate “correlation” with “connectedness” in the graph, and independence with “separation”.

Rule 1: $X$ and $Y$ are d-separated if there is no active path between them.

By “active path”, we mean a path that can be traced without traversing a collider.
Again, a collider is any node where two arrowheads meet. If no active path exists between $X$ and $Y$ then we say that $X$ and $Y$ are d-separated, and the model implies that $X$ and $Y$ are independent.

When we measure a set $Z$ of variables, and take their values as given, the partial correlations of the remaining variables changes character; some correlated variables become uncorrelated, and some uncorrelated variables become correlated. To represent this dynamic in the graph, we need the notion of “partial d-connectedness” or more concretely, “d-connectedness conditioned on a set $Z$ of measurements”.

**Rule 2:** $X$ and $Y$ are d-connected, conditioned on a set of $Z$ nodes, if there is a collider-free path between $X$ and $Y$ that does not traverse a member of $Z$. If no such path exists, we say that $X$ and $Y$ are d-separated by $Z$ or we say that every path between $X$ and $Y$ is “blocked” by $Z$.

A common example used to show that correlation does not imply causation is the fact that ice cream sales are correlated with drowning deaths. When the weather gets warm people tend to both buy ice cream and play in the water, resulting in both increased ice cream sales and drowning deaths. This causal structure is depicted in Figure 5. Here, we see that Ice Cream Sales and Drownings are d-separated given either Temperature or Water Activities. As a result, if we only consider days with the same temperature and/or the same number of people engaging in water activities then the correlation between Ice Cream Sales and Drownings will vanish.
Rule 3: If a collider is a member of the conditioning set \( Z \), or has a descendant in \( Z \), then the collider no longer blocks any path that traces it.

According to Rule 3, conditioning can unblock a blocked path from \( X \) to \( Y \). This is due to the fact that conditioning on a collider or its descendant opens the flow of information between the parents of the collider. For example, \( X \) and \( Y \) are uncorrelated in Figure 4a. However, conditioning on the collider, \( Z \), correlates \( X \) and \( Y \) giving \( X \not\perp\!\!\!\!\perp Y \mid Z \). This phenomenon is known Berkson’s paradox or “explaining away”. To illustrate, consider the example depicted in Figure 4b. It is well known that higher education often affords one a greater salary. Additionally, studies have shown that height also has a positive impact on one’s salary. Let us assume that there are no other determinants of salary and that Height and Education are uncorrelated. If we observe an individual with a high salary that is also short, our belief that the individual is highly educated increases. As a result, we see that observing Salary correlates Education and Height. Similarly, observing an effect or indicator of salary, say the individual’s Ferrari, also correlates Education and Height.

Berkson’s paradox implies that paths containing colliders can be unblocked by conditioning on colliders or their descendants. Let \( \pi' \) be a path from \( X \) to \( Y \) that traces a collider. If for each collider on the path \( \pi' \), either the collider or a descendant of the collider is in the conditioning set \( Z \) then \( \pi' \) is unblocked given \( Z \). The exception to this rule is if \( Z \) also contains a non-collider along the path \( \pi' \) in which case \( X \) and \( Y \) are still blocked given \( Z \). For example, in Figure 3, the path \( F \to C \leftarrow A \to E \) is unblocked given \( C \) or \( D \). However, it is blocked given \( \{A, C\} \) or \( \{A, D\} \).

The following theorem makes explicit the relationship between partial correlation and \( d \)-separation.

**Theorem 1.** Let \( G \) be the causal graph for structural model over a set of variables \( V \). If \( X \in V \) and \( Y \in V \) are d-separated given a set \( Z \subset V \) in \( G \) then the model implies that \( X \) and \( Y \) are independent given \( Z \).

If \( X \) and \( Y \) are d-connected given \( Z \) then \( X \not\perp\!\!\!\!\perp Y \mid Z \) almost surely but \( X \) and \( Y \) may be dependent given \( Z \) for particular parameterizations. For example, it is possible that the values of the coefficients are such that the unblocked paths between \( X \) and \( Y \) perfectly cancel one another.

We use the graph depicted in Figure 3 as an example to illustrate the rules of \( d \)-separation. In this example, \( F \) is d-separated from \( E \) by \( A \) and \( C \). However, \( C \) is not d-separated from \( E \) by \( A \) and \( D \) since conditioning on \( D \) opens the collider \( C \to D \leftarrow B \). Finally \( C \) is d-separated from \( E \) by conditioning on \( A, D, \) and \( B \).

We conclude this section by noting that \( d \)-separation implies conditional independence in all recursive causal models, parametric or not (Pearl, 2000). In linear models, \( d \)-separation implies conditional independence in non-recursive models, as well as recursive models (Spirtes, 1995). Further, all vanishing partial correlations implied by a SEM can be obtained using \( d \)-separation (Pearl, 2000). Finally, in recursive models with independent error terms, these conditional independences represent all of the model’s testable implications (Geiger and Pearl, 1993).
3 Graphical Identification Criteria

3.1 The Single-door Criterion

The single-door criterion is a necessary and sufficient graphical condition for the identification of structural coefficients using regression. When all observable drivers of $Y$ are included in the regression specification then the single-door criterion graphically characterizes conditional exogeneity\(^4\).

**Theorem 2.** (Pearl, 2000) (Single-door Criterion) Let $G$ be any graph for a linear model in which $\beta_0$ is the structural coefficient associated with link $D \rightarrow Y$, and let $G_{\beta_0}$ denote the graph that results when the arrow from $D$ to $Y$ is deleted from $G$. The coefficient $\beta_0$ is identifiable if there exists a set of variables $Z$ such that (i) $Z$ contains no descendant of $Y$ and (ii) $Z$ d-separates $D$ from $Y$ in $G_{\beta_0}$. If $Z$ satisfies these two conditions, then $\beta_0$ is equal to the regression coefficient $r_{YD.Z}$, and we say that $Z$ is a single-door admissible with respect to $\beta_0$. Conversely, if $Z$ does not satisfy these conditions, then $r_{YD.Z}$ is not a consistent estimand of $\beta_0$ (except in rare instances of measure zero).

Consider Figure 6a. As an observable driver of $Y$, $W$ is generally included in the conditioning set when estimating $\beta_0$ using regression. The single-door criterion confirms that, in this case, exogeneity holds and $r_{YD.W} = \beta_0$ since $W$ blocks the spurious path $D \leftarrow Z \rightarrow W \rightarrow Y$ and $D$ is d-separated from $Y$ by $W$ in Figure 6b. Theorem 2 tells us, however, that $Z$ can also be used for adjustment since $Z$ also d-separates $D$ from $Y$ in Figure 6b. Consider, however, Figure 6c. $Z$ and $\{Z,W\}$ satisfy the single-door criterion but $W$ does not. Being a collider, $W$ unblocks the spurious path, $D \leftarrow Z \rightarrow W \leftrightarrow Y$, in violation of Theorem 2, leading to bias if adjusted for. In conclusion, $\beta_0$ is equal to $r_{YD.Z}$ and $r_{YD.WZ}$ in Figures 6a and 6c. However, $\beta_0$ is equal to $r_{YD.W}$ in Figure 6a only.

The intuition for the requirement that $Z$ not be a descendant of $Y$ is depicted in Figures 7a and 7b. We typically do not display the error terms, which can be understood as latent causes. In Figure 7b, we show the error terms explicitly. It

---

\(^4\)Unlike conditional exogeneity, however, the single-door criterion does not require that all observable drivers of $Y$ be included in the regression. This allows us to obtain additional estimands for $\beta_0$, which will prove valuable for robustness testing.
Figure 7: Example showing that adjusting for a descendant of $Y$ induces bias in the estimation of $\beta_0$.

Figure 8: In both graphs, the total effect of $D$ on $Y$ is identifiable even though some of the individual coefficients comprising the effect are not. In (a), the total effect of $D$ on $Y$ is given by $r_{YD}$ while in (b), the total effect of $D$ on $Y$ is given by $r_{YDZ}$.

should now be clear that $Y$ is a collider and conditioning on $Z$ will create spurious correlation between $D$, $U_Y$, and $Y$ leading to bias if adjusted for.

3.2 The Back-door Criterion

In some cases, we may be interested in the total effect of $D$ on $Y$ (given by sums of products of coefficients along all directed paths from $D$ to $Y$) rather than the direct effect (given by the product of coefficients along a single directed path). Moreover, we may be able to identify the total effect even when the individual coefficients comprising it are not identified$^5$. For example, consider Figure 8a. The coefficient associated with the arrow, $W_2 \rightarrow Y$ is not identified. Nevertheless, the total effect of $D$ on $Y$ is $r_{YD}$. The back-door criterion, given below, is a necessary and sufficient condition for the identification of a total effect using regression.

Theorem 3. (Pearl, 2000) (Back-door Criterion) For any two variables $D$ and $Y$ in a model with causal diagram $G$, the total of effect of $D$ on $Y$ is identifiable by regression if and only if there exists a set of measurements $Z$ such that

$^5$The ability to answer policy questions even when individual parameters are not identified was first noted by Marschak (1942). This principle was dubbed “Marschak’s Maxim” by Heckman (2000).
(i) no member of $Z$ is a descendant of $D$; and
(ii) $Z$ $d$-separates $D$ from $Y$ in the subgraph $G_D$ formed by deleting from $G$ all arrows emanating from $D$.

Moreover, if the two conditions are satisfied, then the total effect of $D$ on $Y$ is given by $r_{YDZ}$ and we say that $Z$ is a back-door admissible set.

Returning to the example in Figure 8a we see that the total of effect of $D$ on $Y$ is $r_{YD}$ since $D$ is $d$-separated from $Y$ when arrows leaving $D$ are removed from the graph. Likewise, the total effect of $D$ on $Y$ in Figure 8b is $r_{YDZ}$ since $Z$ blocks all paths between $D$ and $Y$ when arrows emanating from $D$ are removed.

4 Robustness Tests

In this section, we provide a criterion to discern when a shift in the coefficient $r_0$ is indicative of endogeneity bias and when it is not. Additionally, we demonstrate how and when robustness tests are able to detect both observable and unobservable omitted variables.

4.1 Shift-Producing Regressors

Once we have determined that a particular regression coefficient, $r_{YD,W_1}$, identifies $\beta_0$, we may wish to conduct robustness tests in order to check the model assumptions that imply $r_{YD,W_1} = \beta_0$. However, a shift in $r_0$ due to the addition of covariates is not necessarily indicative of model misspecification. In some cases, the model may imply that adding a particular covariate, $W_2$, will induce such a shift in $r_0$. As a result, the shift should be expected and is not indicative of endogeneity bias. We will call such covariates shift-producing. The next theorem uses the model graph to identify regressors that would be shift-producing for almost all parameterizations of the model.

**Theorem 4.** A set of regressors, $W_2$ is shift-producing if $W_1 \cup W_2$ does not satisfy the single-door criterion with respect to $\beta_0$. Conversely, a set of regressors, $W_2$, is not shift-producing if $W_1 \cup W_2$ satisfies the single-door criterion with respect to $\beta_0$.

The single-door criterion gives a necessary and sufficient condition for identification using regression. Therefore, if $W_1 \cup W_2$ satisfies the single-door criterion then $r_{YD,W_1,W_2} = \beta_0 = r_{YD,W_1}$ and adding $W_2$ to the regression of $Y$ on $D$ and $W_1$ should not shift $r_0$. Conversely, we know that defiance of the single-door criterion implies that $r_{YD,W_1,W_2} \neq \beta_0 = r_{YD,W_1}$ (for almost all parameterizations of the model) and, therefore, adding $W_2$ will shift $r_0$ even when $r_{YD,W_1}$ gives an unbiased estimate of $\beta_0$.

Lu and White (2014) indeed identified some shift-producing regressors by noting that variables “driven by $D$” (i.e. descendants of $D$) should not be included in robustness tests for $\beta_0$. This precaution is often justified, but may be overly restrictive.

---

6Descendants of $D$ can be included in the conditioning set, $Z$, so long as spurious paths between $D$ and $Y$ opened by descendants of $D$ are blocked by $Z$. However, such variables can always be excluded without inducing bias.

7We use almost surely in the same sense that we did in the section on $d$-separation: to exclude coincidental cancellations.
Graphical analysis allows us to see when variables driven by $D$ are problematic. Effects of $D$ may be colliders, as in Figure 9 or descendants of $Y$, as in Figure 7a. In Figure 9, conditioning on the collider, $Z$, opens the spurious path, $D \rightarrow Z \leftarrow Y$. Therefore, $Z$ is not single-door admissible and $r_{YD,Z} \neq \beta_0$. Similarly, $Z$ is not single-door admissible in Figure 7a since it is a descendant of $Y$. As we mentioned in Section 3.1, conditioning on $Z$ violates exogeneity by inducing correlation between $D$ and $U_Y$, since $Z$ is a descendant of the collider, $Y$, as shown in Figure 7b. As a result, in both cases, $r_{YD} = \beta_0$ but $r_{YD,Z} \neq \beta_0$.

However, indiscriminately discounting descendants of $D$ is a flawed strategy. In some cases, descendants of $D$ are necessary for identification and can be useful for robustness testing. For example, if our model is the one depicted in Figure 10a, $W_1$ or $W_2$ must be included in the regression in order to obtain an unbiased estimate of $\beta_0$. In Figure 10b, both $W_1$ and $W_2$ must be added to the regression in order to obtain an unbiased estimate of $\beta_0$. Moreover, in the subsequent section on informative regressors, we will see that adding $W_2$ to the regression of $Y$ on $D$ and $W_1$ in Figure 10a constitutes an informative robustness test, demonstrating that descendants of $D$ can also be used for robustness testing.

Finally, discounting variables driven by $D$ is not sufficient to exclude all shift-producing regressors as non-descendants of $D$ may also be shift-producing. For example, in Figure 11, $\beta_0 = r_{YD}$ but $\beta_0 \neq r_{YD,W_1}$ even though $W_1$ is not a descendant of $D$ since conditioning on $W_1$ opens the spurious path, $D \leftrightarrow W_1 \leftrightarrow Y$.

In conclusion, while variables driven by $D$ may be shift-producing, excluding all such variables for identification or robustness testing is both overly restrictive and insufficient for the exclusion of shift-producing regressors. Instead, Theorem 4 provides
4.2 Informative Regressors

A robustness test is informative when a shift in the regression coefficient, \( r_0 \), is indicative of bias, and simultaneously, stability of the coefficient provides evidence that \( r_0 = \beta_0 \). Such tests are characterized by two properties:

(i) The model must imply that \( r_0 \) equals \( \beta_0 \) before and after the addition of regressors.

(ii) There must be the possibility that the robustness test will fail in the case of misspecification\(^8\).

We know from Theorem 2 that the first condition is satisfied if and only if both \( W_1 \) and \( W_1 \cup W_2 \) are single-door admissible. Additionally, if the model is incorrect and \( W_1 \) is not actually single-door admissible but \( W_1 \cup W_2 \) is, then the robustness test will fail. This is the case whenever the model implies that controlling for \( W_1 \) is enough to obtain an unbiased estimate of \( \beta_0 \) but in reality \( W_2 \) is necessary as well. For example, suppose that our model is the one depicted in Figure 10a but, in fact, we are missing an arrow from \( W_2 \) to \( Y \), as in Figure 10b. Then comparing \( r_{Y|D,W_1} \) and \( r_{Y|D,W_1W_2} \) will detect this misspecification since \( r_{Y|D,W_1} \neq r_{Y|D,W_1W_2} \). As a result, we obtain the following theorem:

**Theorem 5.** The robustness test comparing \( r_{Y|D,W_1} \) and \( r_{Y|D,W_1W_2} \) is informative if and only if both \( W_1 \) and \( W_1 \cup W_2 \) are single-door admissible with respect to the coefficient being estimated, \( \beta_0 \).

The above example, depicted by Figures 10a and 10b, demonstrates that informative robustness tests are able to detect misspecification when an observable variable, originally thought to be irrelevant, is, in fact, necessary for an unbiased estimate of \( \beta_0 \).

---

\(^8\)After identifying a set \( W_1 \) that satisfies the conditional exogeneity assumption, Lu and White (2014) advocate adding variables \( W_2 \) such that \( D \perp W_2 | W_1 \) to the regression of \( Y \) on \( D \) and \( W_1 \) for robustness testing. While \( D \perp W_2 | W_1 \) does imply that \( D \perp U | W_1, W_2 \) when \( D \perp U | W_1 \) (Dawid, 1979) so that the model implies \( r_0 = \beta_0 \) before and after the addition of \( W_2 \), the proposed robustness is non-informative. Pearl and Paz (2014) show that \( D \perp W_2 | W_1 \) (see also Lemma 1 below) implies \( r_{Y|D,W_1} = r_{Y|D,W_1W_2} \), even when conditional exogeneity is violated. As a result, the proposed robustness test will ALWAYS pass.

\(^9\)This example also demonstrates our conclusion from the previous section. In some cases, descendants of \( D \) are not only necessary for identification of \( \beta_0 \) but also useful for robustness testing.
This should not be surprising since adding such a variable to the regression shifts $r_0$ by negating or reducing its bias. However, the power of robustness tests is not limited to the detection of endogeneity due to omitted observable variables. Even omitted unobservable variables can be detected by robustness tests. For example, suppose that our model, given by Figure 12a, is misspecified, and the correct specification is given by Figure 12b. In this case (Figure 12b), conditioning on $W_1$ makes $W_2$ an instrumental variable. Since adding an instrumental variable to the regression of $Y$ on $D$ and $W_1$ in the presence of confounding amplifies the bias of $r_0$ (Bhattacharya and Vogt, 2007; Wooldridge, 2009; Pearl, 2010), the bidirected edge between $D$ and $Y$ ensures that $r_{YD,W_2} \neq r_{YD,W_1W_2}$ and comparing the regressions allows us to detect its presence.

In general, if the omitted unobservable variable is a common cause of $D$ and $Y$, and the added covariate, $W_2$, is a non-shift-producing cause (or proxy of a cause) of $D$, then detection is possible. The lemma below characterizes when $r_{YD,W_1} = r_{YD,W_1W_2}$ and will be used to show how and when omitted common causes, both observable and unobservable, can be detected using robustness tests.

**Lemma 1.** For a variable $S$ and a set of variables $W$, $r_{YD,W} = r_{YD,WS}$ if and only if

(i) $\sigma_{SD,W} = 0$ or

(ii) $\sigma_{YS,DW} = 0$,

where $\sigma_{XY,Z}$ is the covariance of $X$ and $Y$ given $Z$.

**Proof.** The sufficiency of Lemma 1 follows from Theorem 4 of (Pearl and Paz, 2014). We now show its necessity by showing that if $r_{YD,W} = r_{YD,WS}$, then either $\sigma_{SD,W} = 0$ or $\sigma_{YS,DW} = 0$.

In general, $r_{YD,WS} = \frac{r_{YD,W} - r_{YS,W}r_{SD,W}}{1 - r_{SD,W}^2\frac{\sigma_{SW}^2}{\sigma_{SW}^2}}$. Since we assume that $r_{YD,W} = r_{YD,WS}$

---

10 Lemma 1 implies that $r_{YD,W} = r_{YD,WS}$ can be characterized in terms of conditional independences. As a result, an informative robustness test simply checks that certain model-implied conditional independences are satisfied in the data. Since all conditional independences implied by the model can be identified using d-separation, robustness testing would be unnecessary if we check all of the testable implications identified using d-separation. (An efficient algorithm that utilizes the graph to enumerate a set, not necessarily minimal, of conditional independences that imply all others is given by Kang and Tian (2009).) However, some model-implied conditional independences may not be relevant to the conditional exogeneity assumption. Therefore, while robustness tests may fail to address certain testable implications that can be identified using d-separation they have the benefit of testing model assumptions that are used specifically in the identification of $\beta_0$. See (Pearl, 2004) and (Chen et al., 2014) for more on testing such assumptions.
we get:

\[ r_{YD,W} = \frac{r_{YD,W} - r_{YS,W}r_{SD,W}}{1 - r_{SD,W}^2 \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2}} \]

\[ 0 = -r_{YD,W} + r_{YD,W}r_{SD,W}^2 \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} + r_{YD,W} - r_{YS,W}r_{SD,W} \]

\[ = r_{YD,W}r_{SD,W}^2 \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} - r_{YS,W}r_{SD,W} \]

\[ = r_{SD,W} \left( r_{YD,W}r_{SD,W} \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} - r_{YS,W} \right) \]

\[ = r_{SD,W} \left( r_{YS,W} - r_{YD,W}r_{SD,W} \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} \right) \]

From the above, we see that either \( r_{SD,W} = 0 \), which implies that \( \sigma_{SD,W} = 0 \), or \( r_{YS,W} - r_{YD,W}r_{SD,W} \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} = 0 \). We also have:

\[ r_{YS,W} - r_{YD,W}r_{SD,W} \frac{\sigma_{D,W}^2}{\sigma_{S,W}^2} \]

\[ = r_{YS,W} \sigma_{S,W}^2 - r_{YD,W} \sigma_{D,W}^2 \]

\[ = \sigma_{YS,DW} \]

Therefore, if \( r_{YD,W} = r_{YD,WS} \) then either \( \sigma_{SD,W} = 0 \) or \( \sigma_{YS,DW} = 0 \). \( \square \)

The theorem below formalizes the conditions under which a robustness test can detect omitted common cause bias, even when the common cause is not observable (and no observable variable can block its influence).

**Theorem 6.** An informative robustness test comparing \( r_{YD,W_1} \) and \( r_{YD,W_1W_2} \) is able to detect omitted common cause bias if and only if there exists a path from \( W_2 \) to \( D \) ending in an arrow or bidirected edge into \( D \) that is not blocked by \( W_1 \).\(^{11,12}\)

**Proof.** Omitted common cause bias of \( r_{YD,W_1} \) implies that model is misspecified so that there is at least one back-door path between \( D \) and \( Y \) that is not blocked by \( W_1 \). \(^{11}\) The sufficiency of this theorem holds almost surely since the correct model may parameterized in such a way that certain paths cancel one another. \(^{12}\) If the conditions of Theorem 6 (\( W_2 \) is connected to \( D \) via a path ending in an arrow or bidirected edge into \( D \) that is not blocked by \( W_1 \) and \( W_1 \) is single-door admissible) hold in the hypothesized model, then the model implies that \( W_2 \) is an instrumental variable given \( W_1 \) (Brito and Pearl, 2002; Chen and Pearl, 2014). As a result, the Hausman test for endogeneity, where the estimate of \( \beta_0 \) using the instrumental variable, \( r_{YW_2,W_1} \), is compared to the estimate using regression, \( r_{YD,W_1} \), is applicable. In fact, both the Hausman and robustness tests check the same model-implied constraints. The Hausman test checks that \( \frac{r_{YW_2,W_1}}{r_{DW_2,W_1}} = r_{YD,W_1} \). Rearranging terms we get \( r_{YW_2,W_1} - r_{YD,W_1}r_{DW_2,W_1} = 0 \), which is true if and only if \( \sigma_{YW_2,DW_1} = 0 \). Likewise, the robustness test checks that \( r_{YD,W_1} = r_{YD,W_1W_2} \), which according to Lemma 1 also holds if and only if \( \sigma_{YW_2,DW_1} = 0 \). The same constraint could also be obtained using d-separation (see footnote 10).

\(^{11}\) The sufficiency of this theorem holds almost surely since the correct model may parameterized in such a way that certain paths cancel one another.

\(^{12}\) If the conditions of Theorem 6 (\( W_2 \) is connected to \( D \) via a path ending in an arrow or bidirected edge into \( D \) that is not blocked by \( W_1 \) and \( W_1 \) is single-door admissible) hold in the hypothesized model, then the model implies that \( W_2 \) is an instrumental variable given \( W_1 \) (Brito and Pearl, 2002; Chen and Pearl, 2014). As a result, the Hausman test for endogeneity, where the estimate of \( \beta_0 \) using the instrumental variable, \( r_{YW_2,W_1} \), is compared to the estimate using regression, \( r_{YD,W_1} \), is applicable. In fact, both the Hausman and robustness tests check the same model-implied constraints. The Hausman test checks that \( \frac{r_{YW_2,W_1}}{r_{DW_2,W_1}} = r_{YD,W_1} \). Rearranging terms we get \( r_{YW_2,W_1} - r_{YD,W_1}r_{DW_2,W_1} = 0 \), which is true if and only if \( \sigma_{YW_2,DW_1} = 0 \). Likewise, the robustness test checks that \( r_{YD,W_1} = r_{YD,W_1W_2} \), which according to Lemma 1 also holds if and only if \( \sigma_{YW_2,DW_1} = 0 \). The same constraint could also be obtained using d-separation (see footnote 10).
In this case, we have a model $M$, which is our misspecified model, and a model $M'$, which is the correct model. $M$ and $M'$ are the same except $M'$ contains at least one back-door path between $D$ and $Y$ that is not blocked by $W_1$ that $M$ does not.

First, we show the sufficiency of Theorem 6 by showing that if there exists a path from $W_2$ to $D$ that is not blocked by $W_1$ and ending in an arrow into $D$ then $r_{Y,D,W_1} \neq r_{Y,D,W_1,W_2}$ (almost surely). We consider two cases. In the first case, all unaccounted for back-door paths are blocked by $W_2$ (i.e. $r_{Y,D,W_1,W_2} = \beta_0$) and in the second case, at least one unaccounted for back-door path is not blocked by $W_2$ (i.e. at least one variable other than $W_2$, either observable or unobservable, must be added to the regression for an unbiased estimate of $\beta_0$). In the first case, $r_{Y,D,W_1} \neq \beta_0$ but $r_{Y,D,W_1,W_2} = \beta_0$. As a result, $r_{Y,D,W_1} \neq r_{Y,D,W_1,W_2}$.

In the second case, there is at least one back-door path that is not blocked by $W_1$ and/or $W_2$. We will show that $r_{Y,D,W_1} \neq r_{Y,D,W_1,W_2}$ by showing that the two conditions of Lemma 1 are violated. First, there exists a path from $W_2$ to $D$ that is not blocked by $W_1$ by assumption and (i) is violated. Second, since there is a path from $W_2$ to $D$ ending in an arrow or bidirected edge into $D$ that is not blocked by $W_1$ and a back-door path between $D$ and $Y$ that is not blocked by $W_1$ or $W_2$ we know that $W_2$ is connected to $Y$ given $D$ and $W_1$ through a path $W_2 \rightarrow D \leftarrow \ldots \rightarrow Y$ and (ii) is violated. As a result, $r_{Y,D,W_1} \neq r_{Y,D,W_1,W_2}$.

We now prove the necessity of Theorem 6. According to Lemma 1, if $r_{Y,D,W_1} \neq r_{Y,D,W_1,W_2}$ then $\sigma_{W_2,D,W_1} \neq 0$ and $\sigma_{W_2,Y,D,W_1} \neq 0$. $\sigma_{W_2,D,W_1} \neq 0$ tells us that there exists a path from $W_2$ to $D$ that is not blocked by $W_1$. We need to show that this path ends in an arrow or bidirected edge into $D$ in $M'$.

Again, we consider two cases. In the first case, at least one unaccounted for back-door path is blocked by $W_2$ and in the second case, none are blocked by $W_2$. If a back-door path is blocked by $W_2$, then there must be a path from $W_2$ to $D$ ending in an arrow or bidirected edge into $D$ in $M'$ and we are done.

In the second case, $\sigma_{W_2,Y,D,W_1} \neq 0$ tells us that there is a path from $W_2$ to $Y$ that is not blocked by $D$ and $W_1$. Since $W_2$ is an informative regressor with respect to $r_{Y,D,W_1}$, $W_1$ and $W_1 \cup \{W_2\}$ are single-door admissible in $M$ and we know that $W_1$ must block any path between $W_2$ and $Y$ in $M$. This implies that $W_1$ blocks all paths between $W_2$ and $Y$ in $M'$ as well, since the only difference between $M'$ and $M$ are back-door paths between $D$ and $Y$ that are not blocked by $W_1$ and/or $W_2$ (we have already considered the case where $W_2$ blocks at least one unaccounted for back-door path). Now, the only possibility of an unblocked path between $W_2$ and $Y$ given $D$ and $W_1$ is that there exists a path from $W_2$ to $D$ ending in either an arrow or bidirected edge into $D$ so that $D$ is a collider, which opens a path $W_2 \rightarrow D \leftarrow \ldots \rightarrow Y$ that goes through an unaccounted for back-door path.

Theorem 6 provides us with the means to answer the question: Have all necessary confounders been accounted for? If we believe that a set $W_1$ is adequate to cover all confounders, we can test this assumption by identifying a non-shift-producing cause (or proxy of a cause) of $D$, whose influence on $D$ is not blocked by $W_1$. We then check the stability of $r_0$ to the addition of this variable.

Lemma 1 and Theorem 6 also demonstrate that some informative robustness tests may be “more informative” than others. For example, suppose that our model is given by Figure 13a but the correct model is the one depicted in Figure 13b. Using
Figure 13: If the true model is given by (b) but the hypothesized model is given by (a) then using $W_1$ for identification of $\beta_0$ and adding $W_2$ to test robustness is able to detect misspecification but not the other way around.

Theorem 6, we see that comparing $r_{YD,W_1}$ and $r_{YD,W_1W_2}$ allows us to detect the latent confounder since $W_2$ is a parent of $D$. However, while both $\{W_2\}$ and $\{W_1, W_2\}$ are single-door admissible, $r_{YD,W_2} = r_{YD,W_1W_2}$, and we cannot detect the latent confounder by comparing these two regressions. This is easily verified using Lemma 1 since $W_1 \bot D | W_2$.

In general, if both $W_1$ and $W_2$ are single-door admissible but $W_2$ blocks all paths between $W_1$ and $D$, adding $W_2$ to the regression of $Y$ on $D$ and $W_1$ may allow the detection of an unobservable confounder but adding $W_1$ to the regression of $Y$ on $D$ and $W_2$ will not. This is formalized in the corollary below.

**Corollary 1.** If both $W_1$ and $W_2$ are single-door admissible but $W_2$ blocks all paths between $W_1$ and $D$, then the robustness test comparing $r_{YD,W_2}$ and $r_{YD,W_1W_2}$ is more informative than the robustness test comparing $r_{YD,W_2}$ and $r_{YD,W_1W_2}$. The former test may be able to detect the presence of an unobservable common cause, while the latter cannot.

**Proof.** In this corollary, we assume that the only difference between the proposed model, $M$, and the correct model, $M'$ is that $M'$ has a back-door path between $D$ and $Y$ not blocked by $W_1$ or $W_2$. As a result, $D \perp W_2 | W_1$ and $r_{YD,W_2} = r_{YD,W_1W_2}$. However, if there is an unblocked path between $W_2$ and $D$ ending in an arrow or bidirected edge into $D$ then Theorem 6 tells us that $r_{YD,W_1} \neq r_{YD,W_1W_2}$. \[\square\]

Moreover, previous work by Kuroki and Miyakawa (2003) and Hahn (2004) have shown that, in this case, $r_{YD,W_2}$ is preferable to $r_{YD,W_1}$ for estimating $\beta_0$ in terms of asymptotic efficiency. The intuition here is that $W_2$ is “closer” to $Y$, hence more effective in reducing variations due to uncontrolled factors.

### 5 Simulations

In order to empirically verify many of our results in Section 4, we generated data using the structures depicted in Figures 13b and 11. For each simulation, we utilize the tr_robust_test function from the testrob Matlab procedure given by Lu and White (2014). This function conducts a Hausman-style test to determine whether a given robustness test passes or fails. In each simulation, we conducted 300 trials with a sample size of 300. Error terms were drawn from Gaussian distributions with mean 0 and
Table 1: Results of simulations derived from the structure depicted in Figure 13b

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>$r_{YD,W_1} = \beta_0$</th>
<th>$r_{YD,W_1} = r_{YD,W_2,W_2}$</th>
<th>$r_{YD,W_2} = \beta_0$</th>
<th>$r_{YD,W_2} = r_{YD,W_1,W_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Trials Passed</td>
<td>0.3%</td>
<td>20.7%</td>
<td>0.3%</td>
<td>93.7%</td>
</tr>
</tbody>
</table>

Table 2: Results of simulations derived from the structure depicted in Figure 11

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>$r_{YD} = \beta_0$</th>
<th>$r_{YD} = r_{YD,W_1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Trials Passed</td>
<td>96.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

variance 1. Additionally, the coefficients for each equation/arrow are drawn uniformly at random from the interval $(0.5, 1)$. We give the percentage of trials for which a given robustness test passes, as well as whether the proposed regression coefficient accurately estimates $\beta_0$, according to a t-test at a 5% significance level.

In the first simulation, we generated data according to Figure 13b. The correlation between the error terms of $D$ and $Y$ is simulated via a latent variable whose values are drawn from a Gaussian distribution with mean 0 and variance 1. In this structure, $\beta_0$ is not identified due to the fact that the error term of $D$ is correlated with the error term of $Y$. Conditioning on either $W_1$, $W_2$, or both does not resolve this issue. In Table 1, we show the results of estimating $\beta_0$ using $r_{YD,W_1}$ and then comparing it to $r_{YD,W_1,W_2}$ as a robustness test. We also show our results from estimating $\beta_0$ using $r_{YD,W_2}$ and comparing it to $r_{YD,W_1,W_2}$ as a robustness test.

The results confirm our conclusions from Section 4. Neither $r_{YD,W_1}$ nor $r_{YD,W_2}$ were able to accurately estimate $\beta_0$ (except in 0.3% of cases). Moreover, in 93.7% of trials, misspecification due to the unaccounted for latent confounder was not detected when adding $W_1$ to the regression of $Y$ on $D$ and $W_2$. In contrast, in 79.3% of trials, misspecification was detected when adding $W_2$, a cause of $D$, to the regression of $Y$ on $D$ and $W_1$.

In our second simulation, we generated data according to Figure 11. The error terms of $D$, $W$, and $Y$ were drawn from a multivariate normal distribution with mean $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and covariance $\begin{pmatrix} 1 & .6 & 0 \\ .6 & 1 & .6 \\ 0 & .6 & 1 \end{pmatrix}$. Here, we verify our finding in Section 4 that even non-descendants of $D$ (i.e. variables not driven by $D$) can be shift-producing. In this case, $r_{YD} = \beta_0$ but adding $W_1$ to the regression shifts the coefficient $r_0$ in 96% of our trials, as shown in the results of our simulation (Table 2).

6 Conclusion

Robustness tests are a valuable tool in testing the structural validity of regression coefficients. However, the covariates involved in such tests must be carefully selected according to the economic model. Certain covariates will produce a shift in $r_0$ when added to the regression, even when it was structurally valid prior to the addition of regressors. A robustness test is informative only when the economic model implies
that the regression coefficient is equal to the structural coefficient before and after the addition of regressors. We have given simple graphical criteria that allow researchers to quickly determine which regression coefficients are structurally valid, according to the model. As a result, they not only aid researchers in the selection of covariates for identification but also for informative robustness tests. Additionally, we have shown that robustness tests are not only able to detect bias due to omitted observable variables but unobservable variables as well. We provided a theorem characterizing when detection of omitted common causes is possible. Finally, we empirically verified many of our findings using simulations.

7 Acknowledgements

We would like thank Angus Deaton, James Heckman, Michael Keane, and Edward Leamer for helpful discussions. This research was supported in parts by grants from NSF #IIS-1302448 and ONR #N00014-10-1-0933 and #N00014-13-1-0153.

References


