# An Introduction to Directed Acyclic Graphs and Markov Equivalencies

Michael J. Throolin

Department of Mathematical Sciences
Montana State University

Date of completion here as May 12, 2023

A writing project submitted in partial fulfillment
of the requirements for the degree

Master of Science in Statistics

# APPROVAL

of a writing project submitted by

Michael J. Throolin

This writing project has been read by the writing project advisor and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Statistics Faculty.

05/15/2023
Date

Katharine Banner
Writing Project Advisor

Date

Andrew Hoegh
Writing Projects Coordinator

**Abstract**

Causal analysis attempts to allow researchers to make causal inference from any type of study, even those without random assignment of some treatment to sample units. A common step in the process of drawing causal inference is the creation of a diagram of a hypothesized causal structure, usually a directed acyclic graph (DAG). When two DAGs yield the same conditional dependencies, they are called Markov equivalent. This equivalence implies that two distinct causal structures may produce similar observable data patterns, making it impossible to determine the true causal relationships between variables from data alone. The use of expert knowledge can help identify the most plausible causal structures from a set of equivalent models in a Markov equivalence class. In this introduction to causal diagrams and equivalence classes, the concept of Markov equivalencies, their importance in causal inference, and two approaches to identifying them will be introduced. Confounding and mediating causal structures will be discussed in detail and methods for simulating data from a confounding structure will be presented. The methods presented in this paper aim to help researchers understand the methodology required to make causal inference from observational data.

# Contents

# 1 Introduction

Statistical analyses seek to understand how variables within a system of data are related. A traditional analysis begins by choosing a statistical model that could approximate a system of interest and then choosing a subset of parameters in the model that could explain aspects of interest in the system. Data are then collected in a manner that allows the researcher to make inference on the chosen parameters. Two types of inference could be sought for, either a causal or associational relationship between variables. An associational relationship implies that, given static external conditions, the relationship between two variables would remain the same. Causal inference seeks to understand how relationships between variables change under dynamic conditions, such as changing the value of a variable in a system (Pearl, 2009). Pearl asserts that, given a large enough sample size, any associational assumption can be tested. "Causal assumptions, in contrast, cannot be verified even in principle, unless one resorts to experimental control (Pearl, 2009)." The researcher should always carefully examine how the data were collected to determine if a causal claim is appropriate or if their results can be generalized beyond the scope of their study.

The gold standard for determining a causal effect is by testing the dynamic conditions between two variables, meaning how changing the value of one covariate effects the other, in a randomized controlled trial. In a randomized controlled trial, treatments are randomly assigned to experimental units and the outcome of interest is measured. The randomizing mechanism should distribute potential confounding variables evenly across treatment levels. Due to the randomizing mechanism being the only prior cause of the treatment, it is safe to assume that the effect of the treatment on the outcome is causal because the potential confounding variables are assumed to be evenly distributed across treatment levels (Huntington-Klein, 2022).

Unfortunately, there are many situations in which a randomized controlled trial

Table 1: Definitions of common terms used in causal inference, adapted from Grace et al. (2012) Huntington-Klein (2022) and Pearl (2000).

| | |
|---|---|
| *Causal Diagram* | A causal diagram consists of a graph where nodes represent different variables and directed edges represent the impact of the variable at the tail of the arrow on the head of the arrow. |
| *Induced dependency* | In a causal model, if an edge connects one node to another, we say the diagram has induced a dependency between the two nodes, as those nodes are no longer independent. |
| *Directed acyclic graph (DAG)* | A directed acyclic graph, such as shown in Figure 1, is a special type of graph where all the edges are directed and there are no cycles (meaning one cannot trace back to the same node if traveling from tail to head along edges) contained in the graph. |
| *Endogenous/Exogenous* | A variable is endogenous if its causes are variables within the structural equation model. A variable is exogenous if its causes are not within the structural equation model. |
| *Markov equivalent* | Two graphs are Markov equivalent if they yield the same induced dependencies. |
| *Parent/Child node* | A node $A$ in a graph is the parent of another node $B$ if there is an arrow pointing from $A$ to $B$. Similarly, $B$ would be called the child of $A$. |
| *Path Coefficient* | The path (or regression) coefficient from A to B represents the causal effect A has on B. |
| *Collider* | If a node along a path only has heads of arrows pointing towards it, such as the node $C$ on the path $A \rightarrow C \leftarrow B$ in Figure 1, it is called a collider. |
| *Skeleton* | The skeleton of a graph is created by replacing all directed edges with undirected edges. |
| *V-Structure* | When two non-adjacent parent nodes have a common child the three variables form a v-structure. |
| *D-separation* | D-separation, short for "dependency separation", is a criterion for determining the conditional independence between two sets of random variables in a causal model given a third set of variables in the same model. |

is not feasible due to the cost, morality, or outright impossibility of performing an experiment. When these situations occur, researchers are left with only observational data, where it is unknown if, in addition to the treatment, other variables have an effect on the outcome. Making causal inference based on observational data is a difficult process that requires making assumptions about how a distribution changes given external conditions. For example, to make causal inference in time series, before after control impact analysis (BACI) assumes that the trajectory of a time series would not change if it were not for an intervention and that no differences exist between a control and intervention other than those which are the result of an intervention (Wauchope et al., 2021). Outside of time series analysis, propensity score matching involves matching pairs of treated and untreated experimental units based on a measure that is assumed to act like randomization would if the data were experimental (Simler-Williamson and Germino, 2022). No matter the analysis, because extra assumptions are required when analyzing observational data, Pearl (2009) states the 'golden rule' of causal inference is that "behind every causal conclusion there must lie some causal assumption that is not testable in observational studies."

When it is not possible to conduct an experiment, structural equation modeling (SEM) is a common methodology used to determine a causal effect. SEM begins by specifying a hypothesized causal structure. Causal structures are often represented as graphs, where a node represents a variable and edges represent the direct relationship between two variables. Edges are represented as arrows that can be either bidirectional, signifying the variables are associated (meaning each node could cause the other), or have a single direction, indicating the root variable has a causal effect on the other. If a graph only has uni-directional arrows as edges, with no paths from one variable back to itself, the diagram is called a directed acyclic graph (DAG). Several common terms that are used in reference to causal structures are

defined in Table 1.

A milestone in linear structural equation modeling was achieved when the method of path coefficients was introduced by Wright (1921) and formalized in Wright (1934). Wright showed how to decompose a structural model into its respective variance-covariance matrix. Linear structural equation models, which Wright decomposed, can be expressed as visual representations of causal assumptions and are powerful tools for finding closed-form expressions for relationships of interest (Pearl, 2013). These tracing rules are outlined in Section 2.2. Pearl (1988) gives a brief history of SEM, describing how Wright's methods were criticized by Niles (1922) and soon forgotten. Fortunately, Wright's rules were rediscovered and used to formalize SEM in the seminal works of Vorobev (1962), Goodman (1970), and Haberman (1974), which connect covariance decomposition with graphs. Causal diagrams have become a common way to visualize the causal assumptions of a structural equation model. Methods to assess the assumptions encoded by a DAG have since been developed, one of the most common being the use of *d-separation*, which is helpful in understanding the concept of Markov equivalence (Pearl, 1988). Definitions of d-separation and Markov equivalence are given in Section 2.1.

In terms of graph-theoretic approaches to structural equation modeling, Grace et al. (2012) provided some practical guidelines for ecologists to follow in their analyses. These guidelines consist of a series of steps, beginning with the development of a hypothesis, structural equation metamodel, and drawing of a causal diagram. A thorough examination and testing of the assumptions implied through the causal diagram ensues. This involves the specification of the structural equation model and its estimation and evaluation. During the evaluation process, a researcher should consider if the model is over-specified or if important edges in the causal diagram are missing. If deemed appropriate during the evaluation process,

the researcher can simplify the model and iterate through the process of model estimation and evaluation until "model-data consistency is declared." When a final model that does not severely disagree with the data has been found, one can examine the underlying "causal assumptions" that the DAG provides. If the researcher feels those assumptions are reasonable given the additional non-statistical knowledge they hold, they may report any "causal conclusions" they find to finish their analysis. Although Grace et al. (2012) wrote these guidelines for ecologists, this approach to structural equation modeling is useful for most other contexts.

# 2 Methods

Understanding the theory behind path analysis and Markov equivalencies requires a knowledge of Wright's tracing rules. This section of the paper will define Wright's tracing rule (Wright, 1934), d-separation and Markov equivalence (Pearl, 1988), and discuss two common causal structures belonging to the same Markov equivalence class— the confounding and mediating causal structures. Wright's tracing rules will be used to show the Markov equivalence between these two structures mathematically, and an example of two causal structures that are not Markov equivalent will be given.

## 2.1 D-Separation and Markov Equivalence

The concept of *d-separation* was formalized for use in causal structures by Verma, Pearl, and Gieger in the 1980s (Pearl, 1988). Recently, Shipley (2000) has developed a d-separation test based on the criterion presented in the works of Pearl (1988). Pearl (2000) summarizes the d-separation test as a method of testing whether a set of X variables is independent of another set, Y variables, given a third set of Z variables: "The idea is to associate 'dependence' with 'connectedness' (i.e., the

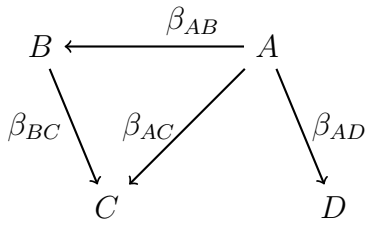


Figure 1: A DAG depicting the causal relationships between four variables, A,B,C, and D, with effects labeled on the edges. A has a causal effect on B,C, and D. B only has an effect on C. C and D do not effect any other variables. This diagram was adapted from Loehlin and Beaujean (2017)

existence of a connecting path) and 'independence' with 'unconnected-ness' (i.e., no path) or 'separation.'"

D-separation is related to the concept of Markov equivalence. When two causal diagrams yield the same sets of conditional independence relationships, they are called *Markov equivalent*. Thus, if two DAGs are Markov equivalent, they will have the same set of d-separation relationships. When classifying a causal diagram into Markov equivalence classes there are two characteristics that need to be considered: the skeleton of the graph and its v-structure. The skeleton of a graph is a copy of the graph with all edges being undirected. A v-structure is formed when two non-adjacent parent nodes have a common child. For example, the only v-structure in Figure 1 is $B \rightarrow C \leftarrow A$. Verma and Pearl (2013) showed that causal graphs are Markov equivalent if and only if they share the same skeleton and v-structures.

## 2.2 Wright's Tracing Rules

Sewall Wright (1934) outlined tracing rules that allows the covariance between two nodes in a causal graph to be computed. There are two concepts that are necessary in the decomposition of covariance: the first being how to find all the causal paths between the variables and the second is determining the root of the causal path. Wright's three tracing rules are:

1. A causal path cannot pass through the same variable twice.

2. Edges in causal paths may be traversed in the opposite direction of the effect they indicate multiple times, but as soon as an edge is traversed in the direction of its causal effect, all subsequent edges in the path must also be traversed in the direction of the effect they indicate.

3. A path may contain a maximum of one bidirectional arrow.

The root of a path is the node which has all arrows pointing away from it. A clear result from the second rule is that these paths are not allowed to contain colliders (i.e, no path contains head-to-head arrows such as $A \rightarrow X \leftarrow B$); in other words, there can only be one root for every path we are tracing while decomposing a DAG. Wright (1934) showed that the covariance between any two nodes in a DAG can be expressed as the sum of the products of a path's coefficients and the variance of its respective roots.

For example, using Wright's tracing rules on the DAG in Figure 1, there are two paths from C to D, namely $\beta_{BC}\beta_{AB}\beta_{AD}$ and $\beta_{AC}\beta_{AD}$. The root of both these paths is A, thus $COV(C, D) = \sigma_A^2(\beta_{BC}\beta_{AB}\beta_{AD} + \beta_{AC}\beta_{AD})$. If we define $\mathbf{X} = \begin{pmatrix} A & B & C & D \end{pmatrix}$, the associated covariance matrix, $Cov(\mathbf{X})$, for the causal diagram in Figure 1 is provided in Equation 1.

$$\begin{pmatrix} \sigma_A^2 & \sigma_A^2\beta_{AB} & \sigma_A^2(\beta_{AC} + \beta_{AB}\beta_{BC}) & \sigma_A^2\beta_{AD} \\ \sigma_A^2\beta_{AB} & \sigma_B^2 & \sigma_B^2\beta_{BC} + \sigma_A^2\beta_{AB}\beta_{AC} & \sigma_A^2\beta_{AB}\beta_{AD} \\ \sigma_A^2(\beta_{AC} + \beta_{AB}\beta_{BC}) & \sigma_B^2\beta_{BC} + \sigma_A^2\beta_{AB}\beta_{AC} & \sigma_C^2 & \sigma_A^2(\beta_{BC}\beta_{AB}\beta_{AD} + \beta_{AC}\beta_{AD}) \\ \sigma_A^2\beta_{AD} & \sigma_A^2\beta_{AB}\beta_{AD} & \sigma_A^2(\beta_{BC}\beta_{AB}\beta_{AD}\beta_{AC}\beta_{AD}) & \sigma_D^2 \end{pmatrix} \quad (1)$$

## 2.3 Confounding and Mediating Causal Structures

Two of the most discussed causal structures in literature are those of confounding and mediating variables. A mediating variable is one that only exists on a graph as

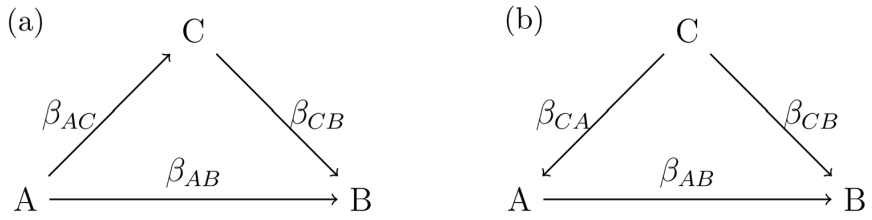


Figure 2: Two DAGs that are markov equivalent. Each edge is labeled with a variable representing the effect a node has on the other. In (a), variable C is a mediator of the effect of A on B. In (b), variable C is confounding the effect of A on B.

a way for one variable to effect another (Huntington-Klein, 2022). For example, Figure 2a shows a causal structure for three variables A, B, and C. Since the only path through C is a way for A to effect B, we can say C is a mediator for A and B.

In the case of of the mediating causal structure shown in Figure 2a, if we let $\mathbf{Y} = \begin{pmatrix} A & B & C \end{pmatrix}$, we can use Wright's Tracing Rules to decompose each path and derive the diagram's respective covariance matrix, provided in 2.

$$Cov(\mathbf{Y}) = \begin{pmatrix} \sigma_A^2 & \sigma_A^2\beta_{AB} + \sigma_A^2\beta_{AC}\beta_{CB} & \sigma_A^2\beta_{AC} \\ \sigma_A^2\beta_{AB} + \sigma_A^2\beta_{AC}\beta_{CB} & \sigma_B^2 & \sigma_A^2\beta_{AB}\beta_{AC} + \sigma_C^2\beta_{CB} \\ \sigma_A^2\beta_{AC} & \sigma_A^2\beta_{AB}\beta_{AC} + \sigma_C^2\beta_{CB} & \sigma_C^2 \end{pmatrix} \tag{2}$$

A confounding variable is one that effects two other variables. For example, Figure 2b shows a causal structure for three variables, A, B, and C. Since C effects both A and B, we can say C confounds the effect of A on B. In the case of the confounding causal structure shown in Figure 2b, if we let $\mathbf{Z} = \begin{pmatrix} A & B & C \end{pmatrix}$, we can use Wright's Tracing Rules to decompose each path and derive the diagram's respective covariance matrix, provided in Equation 3.

$$Cov(\mathbf{Z}) = \begin{pmatrix} \sigma_A^2 & \sigma_A^2\beta_{AB} + \sigma_C^2\beta_{CA}\beta_{CB} & \sigma_C^2\beta_{CA} \\ \sigma_A^2\beta_{AB} + \sigma_C^2\beta_{CA}\beta_{CB} & \sigma_B^2 & \sigma_A^2\beta_{AB}\beta_{CA} + \sigma_C^2\beta_{CB} \\ \sigma_A^2\beta_{AC} & \sigma_A^2\beta_{AB}\beta_{CA} + \sigma_C^2\beta_{CB} & \sigma_C^2 \end{pmatrix} \quad (3)$$

A natural question to ask is 'can a researcher determine, from data alone, whether a variable is confounding or mediating the effect of two other variables?' This is not possible, namely because the covariance between two variables does not determine the unidirectional effect one has on another. In other words, correlation between variables only implies associativity, not causation. This response can be further verified by analyzing the similarities and differences in the covariance matrices of mediating and confounding structures, defined by Equations 2 and 3.

By looking at the specific entries of the confounding and mediating covariance matrices, it can be confirmed that the matrices can only be equal if $\sigma_C^2\beta_{CA} = \sigma_A^2\beta_{AC}$. These quantities, are in fact, equal due to a well-known result in general simple linear regression, where $y = \beta_0 + \beta_{xy}x + \epsilon$, that $\beta_{xy} = \frac{\sigma_y}{\sigma_x}\rho_{xy}$ (Pearl, 2013). Extending this to our question, $\sigma_C^2\beta_{CA} = \sigma_C^2\frac{\sigma_A}{\sigma_C}\rho_{CA} = \sigma_C\sigma_A\rho_{AC} = \sigma_A^2\frac{\sigma_C}{\sigma_A}\rho_{AC} = \sigma_A^2\beta_{AC}$.

Therefore, the covariance matrices induced by confounding and mediating causal structures are equivalent. This means it would be impossible to distinguish the two structures given only raw data as they generate the same induced dependencies; in other words, the confounding and mediating causal structures are Markov equivalent. The theorem proved by Verma and Pearl (2013) makes it much easier to determine that these causal structures are equivalent as both the confounding and mediating causal structures, drawn in Figure 2, share the same skeleton and have no v-structures.

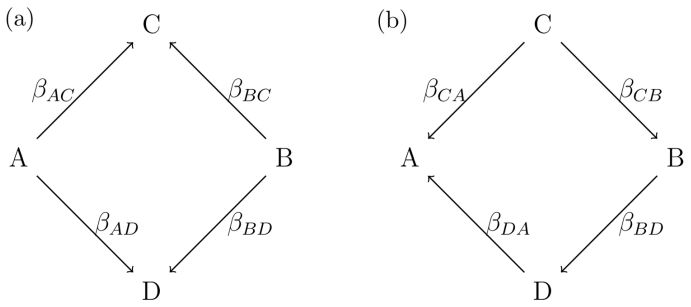


Figure 3: Two DAGs that share the same skeleton but are not Markov equivalent as their v-structures are different. (a) has two v-structures, $A \to C \leftarrow B$ and $A \to D \leftarrow B$. (b) has a single v-structure, $C \to A \leftarrow D$.

### 2.4 Non Markov-Equivalent Structures

An example of two DAGs that are not Markov equivalent is shown in Figure 3. Both DAGs pictured in this figure have the same skeleton. However, Figure 3a has two v-structures, $A \to C \leftarrow B$ and $A \to D \leftarrow B$, and 3b only has one, $C \to A \leftarrow D$. Thus, these two diagrams are not Markov equivalent and have different induced dependencies. For example, in Figure 3a, using Wright's tracing rules it is clear that A and B are independent, or $COV(A, B) = 0$. However, in Figure 3b no variables are independent of each other. Specifically, $COV(A, B) = \sigma_B^2 \beta_{BD} \beta_{DA} + \sigma_C^2 \beta_{CA} \beta_{CB}$, which is non-zero whenever $\sigma_A^2 > 0$ or $\sigma_C^2 > 0$.

## 3 Simulation Study

Fitting models to simulated data is common practice in many fields of statistics to verify that models are functioning correctly. In this section data will be simulated from a confounding causal structure and the underlying parameters of the structural equation model will be estimated. To simulate data, one must follow a series of four steps, outlined below.

1. *Choose a causal diagram to simulate.* The first step of simulating data is to choose a causal diagram. For the purposes of this tutorial, the confounding

structure shown in Figure 2b will be simulated.

2. *Specify the model.* The model must follow the assumptions encoded in the causal diagram. Each node may follow a different distribution. In this simulation the model is specified as follows:

$$\begin{aligned} C &\overset{iid}{\sim} N(\mu_C, \sigma_C^2), \\ A &\overset{iid}{\sim} N(\beta_{0_A} + \beta_{CA} C, \epsilon_A^2), \\ B &\overset{iid}{\sim} N(\beta_{0_B} + \beta_{AB} A + \beta_{CB} C, \epsilon_B^2) \end{aligned} \tag{4}$$

3. *Assign data-generating values to all model parameters.* In this simulation the parameters will be assigned the values $\mu_C = 3$, $\epsilon_A = \epsilon_B = \sigma_C = 1$, $\beta_{0_A} = \beta_{0_B} = 0$, $\beta_{CA} = 2$, $\beta_{AB} = 3$, and $\beta_{CB} = 4$.

4. *Generate data from the model described in Step 2.* For this simulation, 100 observations are going to be simulated from Equation 4.

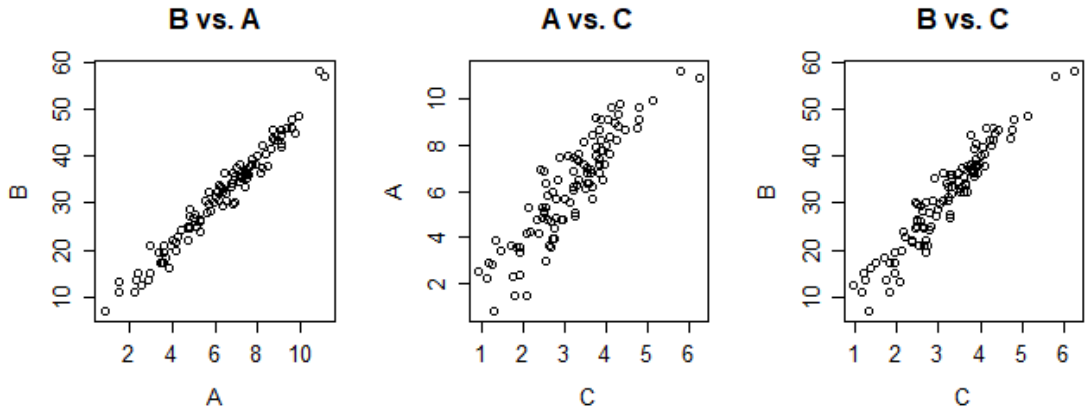


Figure 4: Scatter plots depicting relationships between the variables of A, B, and C. Values were generated from the model specified in Equation 4.

It is good practice to visualize data throughout an analysis. The relationship between the 100 observations simulated from Equation 4 can be observed in the

Table 2: Estimated model parameters for a sample size of 100 observations, generated from Equation 4.

| Parameter | True Value | Point Estimate | 95 % Confidence Interval |
|---|---|---|---|
| $\beta_{AC}$ | 2 | 2.01 | (1.82,2.21) |
| $\beta_{AB}$ | 3 | 2.97 | (2.86,3.21) |
| $\beta_{CB}$ | 4 | 3.92 | (3.30,4.06) |
| $\sigma_C^2$ | 1 | 1.00 | |
| $\epsilon_A$ | 1 | 0.97 | |
| $\epsilon_B$ | 1 | 1.07 | |

pairwise scatter plots shown in Figure 4. Through visual inspection of the relationship between A vs. C, one can verify that $\beta_{0_A} \approx 0$ and $\beta_{CA} \approx 2$.

There are several methods within causal inference available to estimate the parameters of an underlying model. Regression is an appropriate method to estimate parameters when there are no severe violations of the assumptions of linearity, homoscedasticity, normality of residuals, and influential points. An assumption of regression within the context of causal inference that becomes particularly important is that the covariates in the model are exogenous to, or uncorrelated with, the error term Huntington-Klein (2022).

To estimate all the parameters in the assumed structural equation model, two models need to be fit: $A \sim N(\beta_{0_A} + \beta_{CA}C, \epsilon_A^2)$ and $B \sim (\beta_{0_B} + \beta_{AB}A + \beta_{CB}C, \epsilon_B^2)$. The regression coefficients are estimated with their standard errors and the residual standard error from these models are used to estimate $\epsilon_A$ and $\epsilon_B$ respectively. The sample mean and sample variance of C are estimators for $\mu_C$ and $\sigma_C^2$ respectively.

The data visualized in Figure 4 were analyzed using regression in R (R Core Team, 2023). The code written to estimate the parameters in the underlying model is provided in the appendix (Section 6.1). The estimates, along with their associated 95% confidence intervals for the data visualized in Figure 4 are listed in Table 2. The process of simulation and estimation was iterated 1000 times and plots of the

parameter estimates from these thousand iterations are shown in Figure 5. These plots show that the parameter estimates were all centered around their true data-generating values.

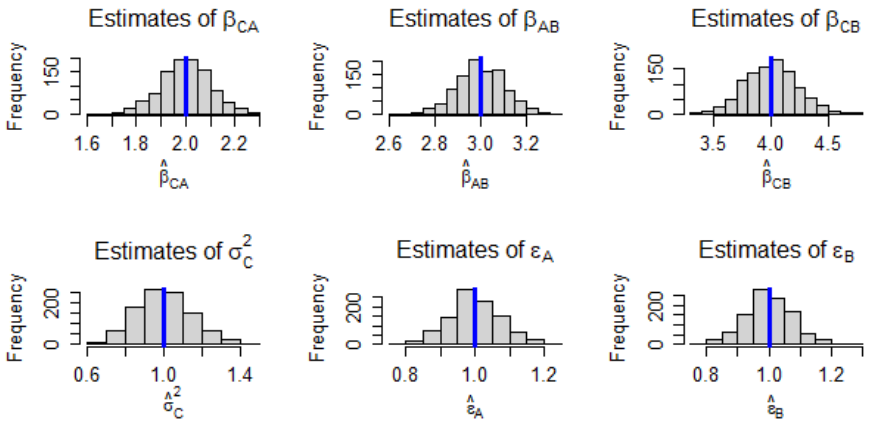


Figure 5: Histograms of parameter estimates, using regression, from 1000 iterations of data simulated from the model specified in Equation 4. The data-generating values are plotted as vertical lines in each plot.

# 4 Conclusion

There has been much discussion about Markov equivalence classes and the issues researchers face when trying to distinguish causal structures from observational data. When doing a causal analysis, unless the researcher has outside information, their research can at most only determine which Markov equivalence class the underlying causal structure lies in. The simulation study presented in Section 3 gave an example of how to conduct a cursory path analysis of a hypothesized structure.

Several R packages exist for more in-depth analysis in structural equation modeling such as `dagR` (Breitling et al., 2022), `DAGitty` (Textor et al., 2016) and `lavaan` (Rosseel, 2012). `DagR` is a visualization tool that integrates well with a variety of other R packages. `DAGitty` is used to construct DAGs to find underlying d-separation relationships. As a visualization tool, `DAGitty` is particularly adept at

identifying confounding variables. `Lavaan` can simulate data from a specified DAG. It also can analyze data through path analysis, confirmatory factor analysis, or latent variable modeling. However a researcher chooses to analyze data, when reporting the results of a causal analysis it is important to remember Pearl's golden rule that "behind every causal conclusion there must lie some causal assumption that is not testable in observational studies" and state all causal assumptions they made during their analysis (Pearl, 2009).

# 5 Acknowledgements

I would like to thank Dr. Kathryn Irvine and William Hammond for holding a weekly reading group in causal inference with me. Dr. Irvine in particular helped improve my understanding of causal inference and without her I would not have been able to write this paper. I would also like to thank Katharine Banner for her feedback as this paper developed.

# 6 Appendix

## 6.1 Simulation Code

```
# Function that simulates data from a confounding variable
    sim_confound <- \(n = 100, beta_ca = 2, beta_ab =3, beta_cb = 4,
                      mu_c = 3, eps_A = 1,
                      eps_B = 1,sigma_C = 1){
      C = rnorm(n, mean = mu_c, sd = eps_C^2)
      A = beta_ca*c +rnorm(n, mean = 0, sd = eps_A^2)
      B = beta_ab*a + beta_cb*c+rnorm(n, mean = 0, sd = eps_B^2)
      return(data.frame(A,B,C))
    }
#Number of simulations
    n_sims<-1000
#Initialize values
    beta_ca <- rep(NA, n_sims)
    beta_ab <- rep(NA, n_sims)
    beta_cb <- rep(NA, n_sims)
    varC <- rep(NA, n_sims)
    epsA <- rep(NA, n_sims)
    epsB <- rep(NA, n_sims)


# Iterate
    for(i in 1:n_sims){
      dat <- sim_confound()
```

```
## Estimate path coefficients via regression
  beta_ca[i] <- lm(A~C, data = dat)$coefficients["C"]
  beta_ab[i] <- lm(B~A+C, data = dat)$coefficients["A"]
  beta_cb[i] <- lm(B~C+A, data= dat)$coefficients["C"]

## Estimate epsilons
  varC[i] <- var(dat$C)
  epsA[i] <- sigma(lm(A~C, data = dat))
  epsB[i]<- sigma(lm(B~C+A, data= dat))
}
```

# 7 References


Breitling, L. P., Duan, C., Dragomir, A. D., and Luta, G. (2022). Using dagr to identify minimal sufficient adjustment sets and to simulate data based on directed acyclic graphs. *International Journal of Epidemiology*.

Goodman, L. A. (1970). The multivariate analysis of qualitative data: Interaction among multiple classifications. *Journal of the American Statistical Association*, 65(330):226–256.

Grace, J. B., Schoolmaster Jr, D. R., Guntenspergen, G. R., Little, A. M., Mitchell, B. R., Miller, K. M., and Schweiger, E. W. (2012). Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere*, 3(8):1–44.

Haberman, S. J. (1974). *The General Log-Linear Model*. Phd thesis, University of Chicago.

Huntington-Klein, N. (2022). *The Effect*. Chapman & Hall.

Loehlin, J. C. and Beaujean, A. A. (2017). *Latent variable models: an introduction to factor, path, and structural equation analysis*. Routledge/ Taylor & Francis Group, New York, fifth edition.

Niles, H. E. (1922). Correlation, causation, and wright theory of "path coefficients.". *Genetics*, 7(3):258–273.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems : networks of plausible inference*, chapter 3. Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann Publishers, Inc., San Francisco, California, revised second edition.

Pearl, J. (2000). *Causality : models, reasoning, and inference*, chapter 11.1.2, pages "335–337". Cambridge University Press, Cambridge, second edition.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146. Publisher: Amer. Statist. Assoc., the Bernoulli Soc., the Inst. Math. Statist., and the Statist. Soc. Canada.

Pearl, J. (2013). Linear Models: A Useful "Microscope" for Causal Analysis. *Journal of Causal Inference*, 1(1):155–170.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2):1–36.

Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling: A Multidisciplinary Journal*, 7(2):206–218.

Simler-Williamson, A. B. and Germino, M. J. (2022). Statistical considerations of nonrandom treatment applications reveal region-wide benefits of widespread post-fire restoration action. *Nature Communications*, 13(1):3472.

Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., and Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the r package 'dagitty'. *International Journal of Epidemiology*, 45(6):1887–1894.

Verma, T. S. and Pearl, J. (2013). On the Equivalence of Causal Models.

Vorobev, N. N. (1962). Consistent families of measures and their extensions. *Theory of Probability and its Applications*, 7(2):147–163.

Wauchope, H. S., Amano, T., Geldmann, J., Johnston, A., Simmons, B. I., Sutherland, W. J., and Jones, J. P. (2021). Evaluating Impact Using Time-Series Data. *Trends in Ecology & Evolution*, 36(3):196–205. Publisher: Elsevier.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20(7):557–585.

Wright, S. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics*, 5(3):161–215. Publisher: Institute of Mathematical Statistics.