

Backdoor Criterion

Definition (The Backdoor Criterion)

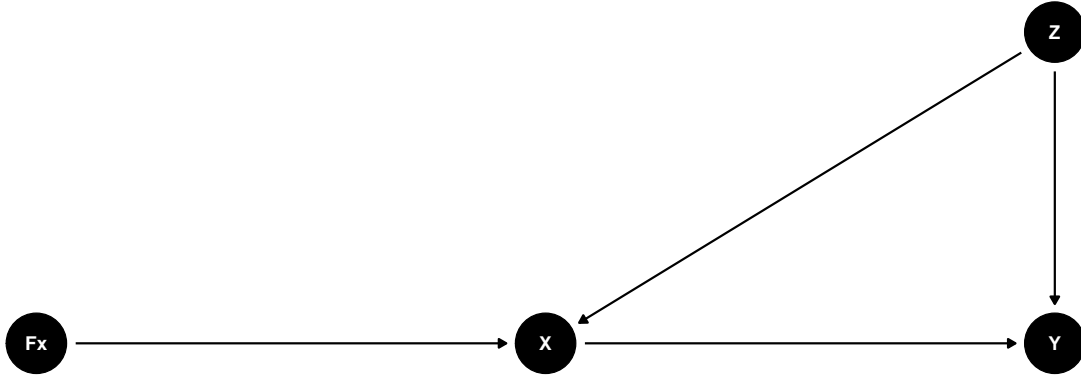
Given an ordered pair of variables (X, Y) in a directed acyclic graph G , a set of variables Z satisfies the backdoor criterion relative to (X, Y) if no node in Z is a descendant of X , and Z blocks every path between X and Y that contains an arrow into X .

Causal Effect

If a set of variables Z satisfies the backdoor criterion for X and Y , then the causal effect of X on Y is given by the formula:

$$P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, Z = z)P(Z = z)$$

Introducing regime indicator (F_X) which takes values in $\{do(x'), \text{idle (no intervention)}\}$ for proof and an example of graphical model, G' . Note that a graphical model without F_X is G and $G' = G \cup \{F_i \rightarrow X_i\}$.



The conditional probability, $P(x_i \mid pa(x_i, G'))$, introduces the new parent set of X_i , $PA'_i = PA_i \cup \{F_i\}$

$$P(x_i \mid pa(x_i, G')) = \begin{cases} P(x_i \mid pa(x_i, G)) & \text{if } F_i = \text{idle} \\ 1 & \text{if } F_i = do(x'_i) \text{ and } x_i = x'_i \\ 0 & \text{if } F_i = do(x'_i) \text{ and } x_i \neq x'_i \end{cases}$$

Now, we are ready for proof!

Proof for $P(y \mid do(x)) = \sum_z P(y \mid z, x)P(z)$:

$$\begin{aligned}
& P(y \mid do(x)) \\
&= P(y \mid F_x = do(x)) \\
&= \sum_z P(y, z \mid F_x = do(x)) \\
&= \sum_z P(y \mid z, F_x = do(x))P(z \mid F_x = do(x)) \\
&= \sum_z P(y \mid z, x, F_x = do(x))P(z \mid F_x = do(x)) \\
&=^1 \sum_z P(y \mid z, x, F_x = do(x))P(z) \\
&=^2 \sum_z P(y \mid z, x)P(z)
\end{aligned}$$

- (1) Parental Markov condition (local Markov condition): A necessary and sufficient condition for a probability distribution P to be Markov relative a DAG G is that every variable be independent of all its nondescendants (in G), conditional on its parents.

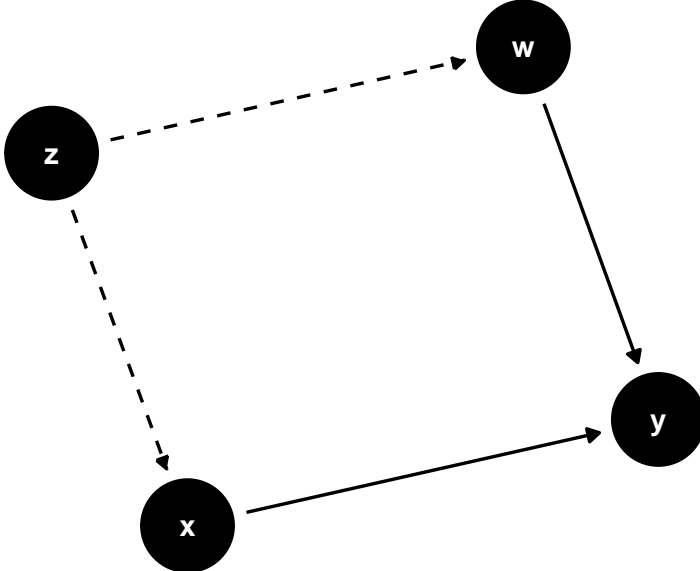
By this theorem, $F_x \perp\!\!\!\perp Z$

- (2) By backdoor condition, F_x and Y are d-separated, conditional on Z and $X \Rightarrow F_x \perp\!\!\!\perp Y \mid Z, X$

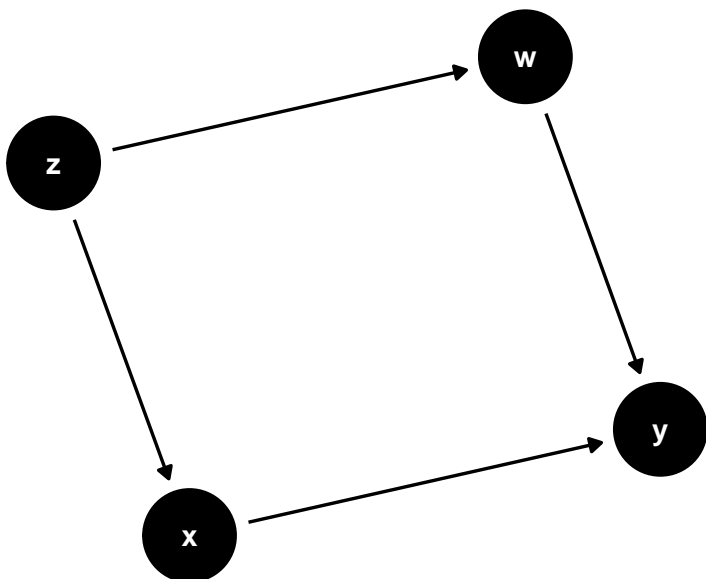
PA(X) always satisfies backdoor criterion:

$$P(y \mid pa(x)) = \sum_z P(y \mid pa(x), x)P(pa(x))$$

Backdoor Criterion Example1:



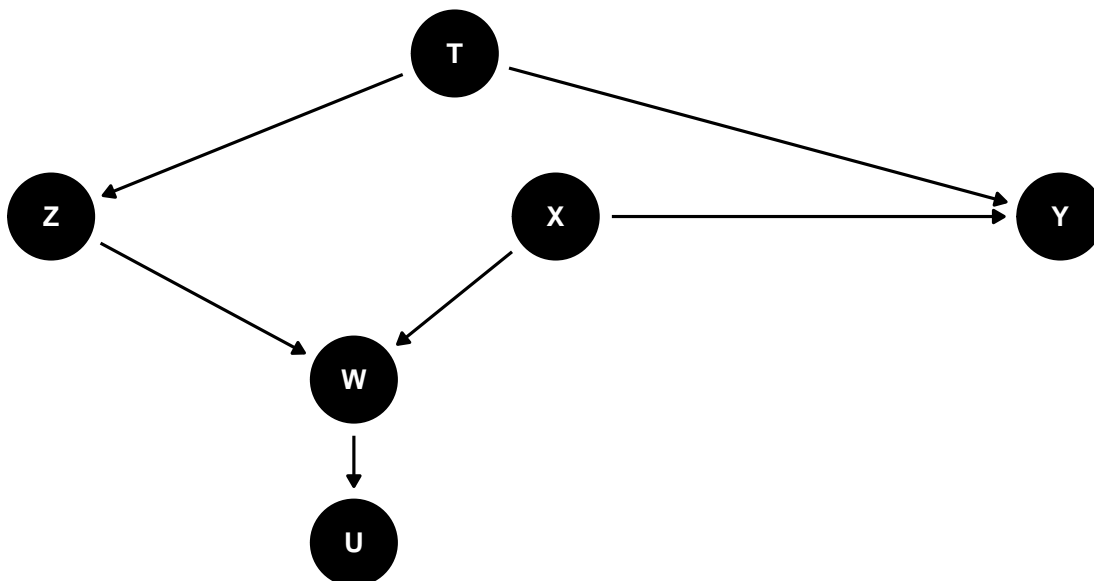
Here, we want to estimate the effect of smoking cigarette (x) on lung function (y). Weight, w , is also measured and we know the type of jobs affect both weight and the choice to smoke cigarette. However, the study did not record the type of jobs. Though, z is not recorded, we can estimate $P(Y = y \mid do(X = x))$ using the backdoor criterion. Here, w is not a descendant of x and also blocks the backdoor path $x \leftarrow z \rightarrow w \rightarrow y$. Thus, $P(Y = y \mid do(X = x)) = \sum_w P(Y = y \mid X = x, W = w)P(W = w)$



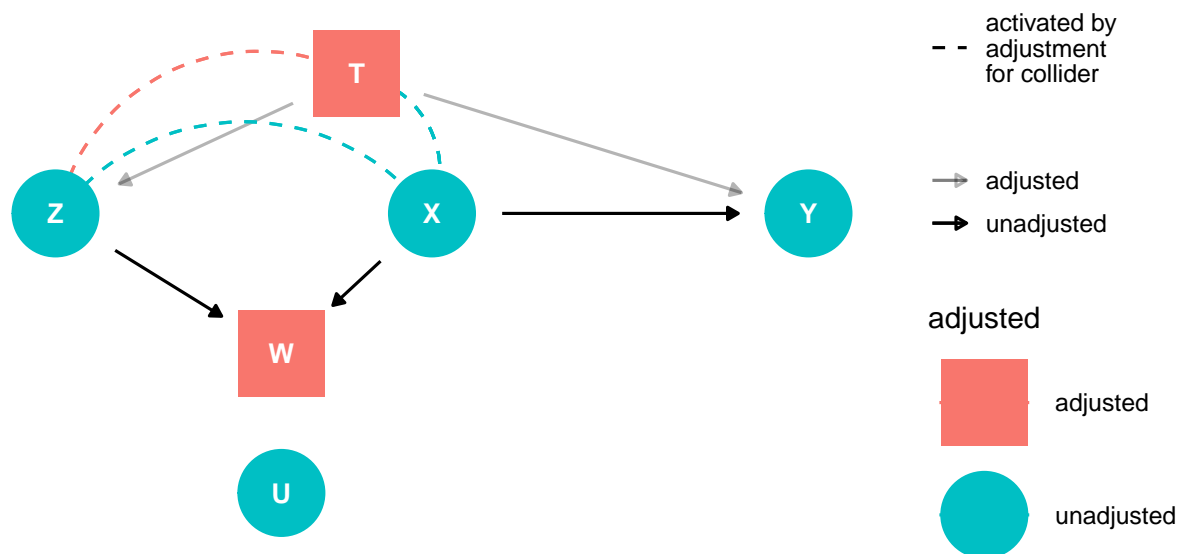
Now assume z is observed as well as w . Then, blocking the backdoor path can be done by adjusting for either z or w . This is useful because the choice can be made depending on which variable is more convenient to measure.

$$P(Y = y \mid do(X = x)) = \sum_w P(Y = y \mid X = x, W = w)P(W = w) = \sum_w P(Y = y \mid X = x, Z = z)P(Z = z)$$

Backdoor Criterion Example2:



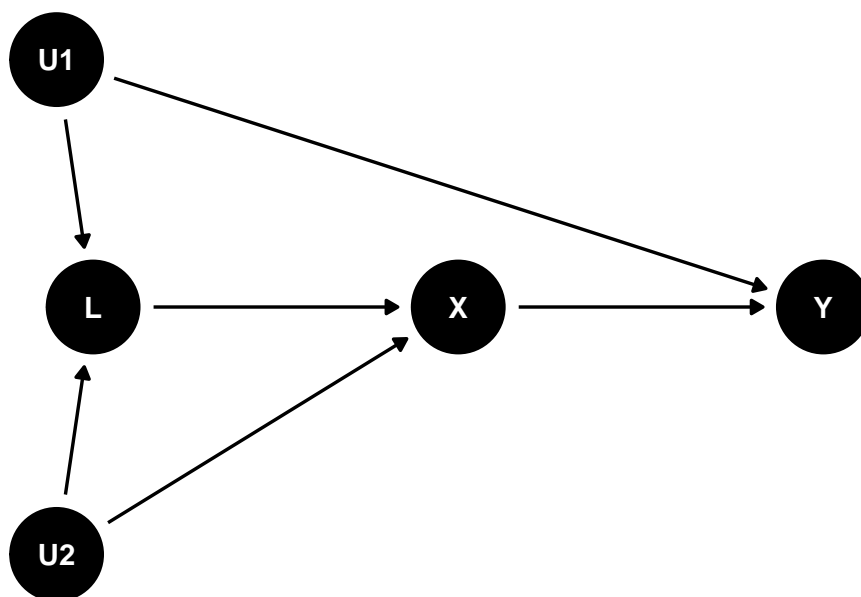
Here, we want to estimate the effect of X on Y for a specific value w of W . Though W is a collider and conditioning on W opens a path, T (not a descendant of X) can be used to block the spurious path $X \rightarrow W \leftarrow Z \leftrightarrow T \rightarrow Y$.



We can see the backdoor path is blocked when adjusting for W and T.

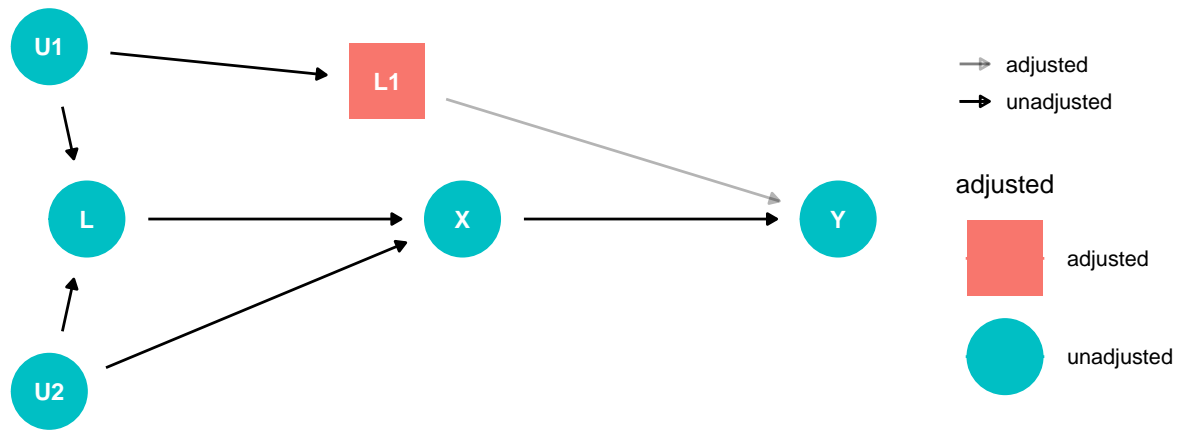
Thus, $P(Y = y \mid do(X = x), W = w) = \sum_t P(Y = y \mid X = x, W = w, T = t)P(T = t \mid W = w)$

Backdoor Criterion Example3:



Here, we want to estimate the effect of X on Y. Conditioning on L would block the backdoor path, $X \leftarrow L \leftarrow U_1 \rightarrow Y$, but would simultaneously open a new backdoor path, $X \rightarrow U_2 \rightarrow L \leftarrow U_1 \rightarrow Y$. Namely, the attempt to block the confounding path brings about a selection bias. Here, A solution would be to measure either (i) a variable L_1 between U_1 and either A or Y, or (ii) a variable L_2 between u_2 and either A or L.

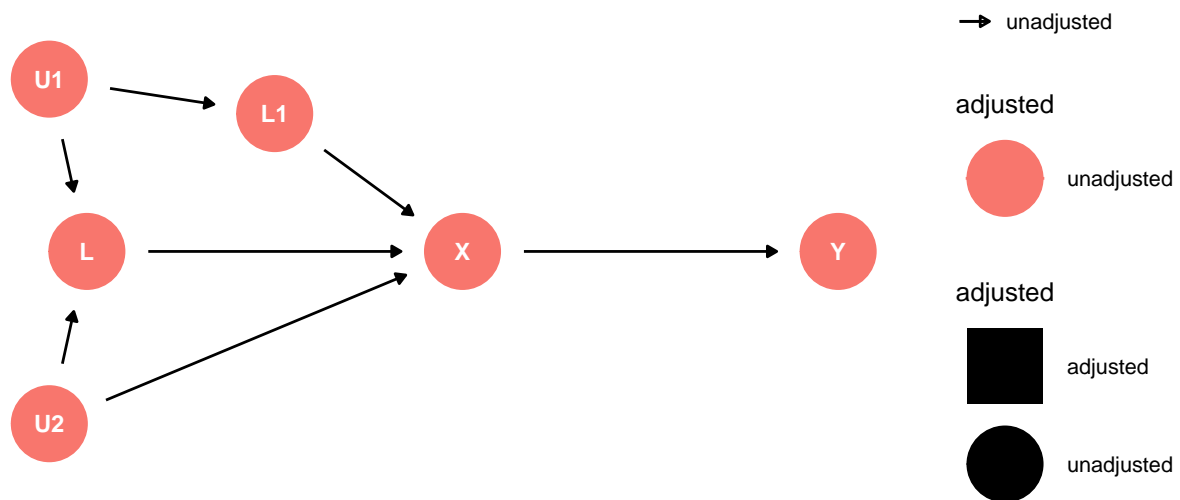
Example3 Solution1 (L_1 between U_1 and Y):



Adjusting for L_1 blocks the backdoor path.

Thus, $P(Y = y \mid do(X = x)) = \sum_{l_1} P(Y = y \mid X = x, L_1 = l_1)P(L_1 = l_1)$

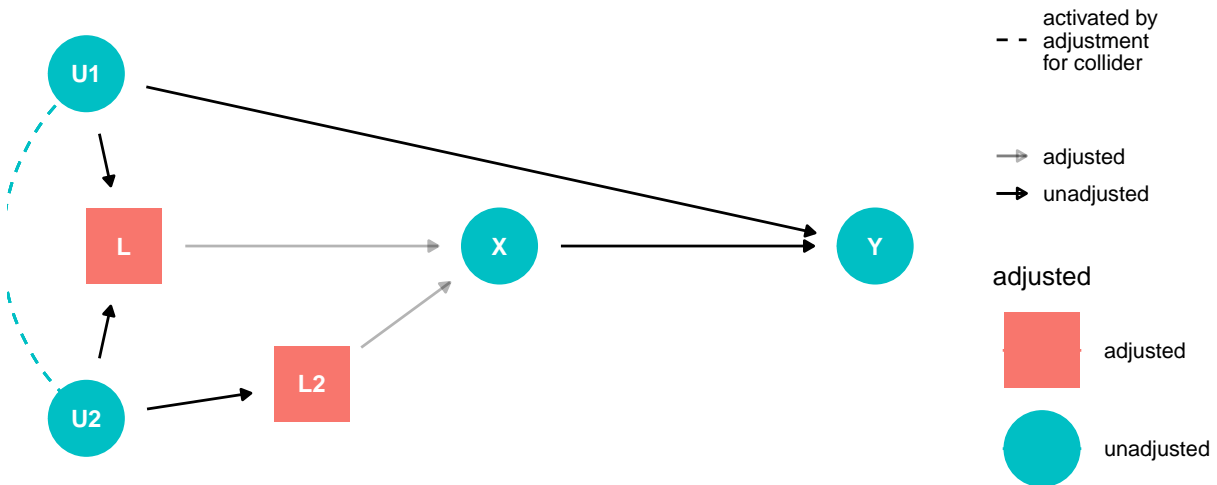
Example3 Solution1 (L_1 between U_1 and X):



There is no backdoor path by adding L_1 variable.

Thus, $P(Y = y \mid do(X = x)) = P(Y = y \mid X = x)$

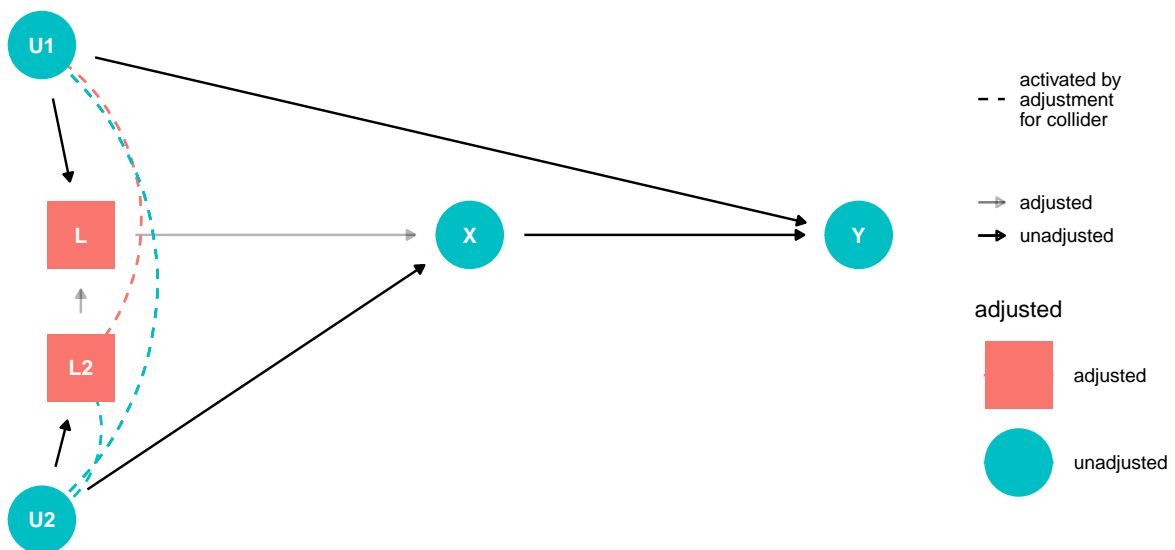
Example3 Solution2 (L_2 between U_2 and X):



Adjusting for L and L_2 blocks the backdoor path.

Thus, $P(Y = y \mid do(X = x)) = \sum_{l, l_2} P(Y = y \mid X = x, L = l, L_2 = l_2)P(L = l, L_2 = l_2)$

Example3 Solution2 (L_2 between U_2 and L):



Adjusting for L and L_2 blocks the backdoor path.

Thus, $P(Y = y \mid do(X = x)) = \sum_{l, l_2} P(Y = y \mid X = x, L = l, L_2 = l_2)P(L = l, L_2 = l_2)$

Reference

Pearl, J., Glymour, M., & Jewell, N. P. (2019). Causal inference in statistics a primer. Wiley.

Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.