## 3.5 Graphical Tests of Identifiability

Figure 3.7 shows simple diagrams in which $P(y|\hat{x})$ cannot be identified owing to the presence of a "bow" pattern—a confounding arc (dashed) embracing a causal link between $X$ and $Y$. A confounding arc represents the existence in the diagram of a back-door path that contains only unobserved variables and has no converging arrows. For example, the path $X,\ Z_0,\ B,\ Z_3$ in Figure 3.1 can be represented as a confounding arc between $X$ and $Z_3$. A bow pattern represents an equation $y = f_Y(x, u, \epsilon_Y)$, where $U$ is unobserved and dependent on $X$. Such an equation does not permit the identification of causal effects, since any portion of the observed dependence between $X$ and $Y$ may always be attributed to spurious dependencies mediated by $U$.

The presence of a bow pattern prevents the identification of $P(y|\hat{x})$ even when it is found in the context of a larger graph, as in Figure 3.7(b). This is in contrast to linear models, where the addition of an arc to a bow pattern can render $P(y|\hat{x})$ identifiable (see Chapter 5, Figure 5.9). For example, if $Y$ is related to $X$ via a linear relation $y = bx + u$, where $U$ is an unobserved disturbance possibly correlated with $X$, then $b = \frac{\partial}{\partial x}E(Y|\hat{x})$ is not identifiable. However, adding an arc $Z \rightarrow X$ to the structure (i.e., finding a variable $Z$ that is correlated with $X$ but not with $U$) would facilitate the computation of $E(Y|\hat{x})$ via the instrumental variable formula (Bowden and Turkington 1984; see also Chapter 5):

$$b \triangleq \frac{\partial}{\partial x}E(Y|\hat{x}) = \frac{E(Y|z)}{E(X|z)} = \frac{r_{YZ}}{r_{XZ}}. \tag{3.48}$$

In nonparametric models, adding an instrumental variable $Z$ to a bow pattern (Figure 3.7(b)) does not permit the identification of $P(y|\hat{x})$. This is a familiar problem in the analysis of clinical trials in which treatment assignment ($Z$) is randomized (hence, no link enters $Z$) but compliance is imperfect (see Chapter 8). The confounding arc between $X$ and $Y$ in Figure 3.7(b) represents unmeasurable factors that influence subjects' choice of treatment ($X$) as well as subjects' response to treatment ($Y$). In such trials, it is not possible to obtain an unbiased estimate of the treatment effect $P(y|\hat{x})$ without making additional assumptions on the nature of the interactions between compliance and response (as is done, for example, in the potential-outcome approach to instrumental variables developed in Imbens and Angrist 1994 and Angrist et al. 1996. Although the added arc $Z \rightarrow X$ permits us to calculate bounds on $P(y|\hat{x})$ (Robins 1989; sec. 1g; Manski 1990; Balke and Pearl 1997) and the upper and lower bounds may even coincide for certain types of distributions $P(x, y, z)$ (Section 8.2.4), there is no way of computing $P(y|\hat{x})$ for *every* positive distribution $P(x, y, z)$, as required by Definition 3.2.4.

In general, the addition of arcs to a causal diagram can impede, but never assist, the identification of causal effects in nonparametric models. This is because such addition reduces the set of $d$-separation conditions carried by the diagram; hence, if a causal effect derivation fails in the original diagram, it is bound to fail in the augmented diagram as well. Conversely, any causal effect derivation that succeeds in the augmented diagram (by a sequence of symbolic transformations, as in Corollary 3.4.2) would succeed in the original diagram.

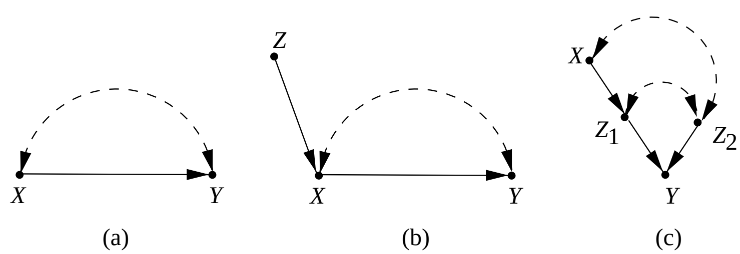


Figure 3.7: (a) A bow pattern: a confounding arc embracing a causal link $X \rightarrow Y$, thus preventing the identification of $P(y|\hat{x})$ even in the presence of an instrumental variable $Z$, as in (b). (c) A bowless graph that still prohibits the identification of $P(y|\hat{x})$.

Our ability to compute $P(y_1|\hat{x})$ and $P(y_2|\hat{x})$ for pairs $(Y_1, Y_2)$ of singleton variables does not ensure our ability to compute joint distributions, such as $P(y_1, y_2|\hat{x})$. Figure 3.7(c), for example, shows a causal diagram where both $P(z_1|\hat{x})$ and $P(z_2|\hat{x})$ are computable yet $P(z_1, z_2|\hat{x})$ is not. Consequently, we cannot compute $P(y|\hat{x})$. It is interesting to note that this diagram is the smallest graph that does not contain a bow pattern and still presents an uncomputable causal effect.

Another interesting feature demonstrated by Figure 3.7(c) is that computing the effect of a joint intervention is often easier than computing the effects of its constituent singleton interventions.[6] Here, it is possible to compute $P(y|\hat{x}, \hat{z}_2)$ and $P(y|\hat{x}, \hat{z}_1)$, yet there is no way of computing $P(y|\hat{x})$. For example, the former can be evaluated by invoking Rule 2 in $G_{\overline{X}\underline{Z_2}}$, giving

$$P(y|\hat{x}, \hat{z}_2) = \sum_{z_1} P(y|z_1, \hat{x}, \hat{z}_2)P(z_1|\hat{x}, \hat{z}_2) = \sum_{z_1} P(y|z_1, x, z_2)P(z_1|x). \tag{3.49}$$

However, Rule 2 cannot be used to convert $P(z_1|\hat{x}, z_2)$ into $P(z_1|x, z_2)$ because, when conditioned on $Z_2$, $X$ and $Z_1$ are $d$-connected in $G_{\underline{X}}$ (through the dashed lines). A general approach to computing the effect of joint interventions is developed in Pearl and Robins (1995); this is described in Chapter 4 (Section 4.4).

### 3.5.1 Identifying Models

Figure 3.8 shows simple diagrams in which the causal effect of $X$ on $Y$ is identifiable (where $X$ and $Y$ are single variables). Such models are called "identifying" because their structures communicate a sufficient number of assumptions (missing links) to permit the identification of the target quantity $P(y|\hat{x})$. Latent variables are not shown explicitly in these diagrams; rather, such variables are implicit in the confounding arcs (dashed). Every causal diagram with latent variables can be converted to an equivalent diagram involving measured variables interconnected by arrows and confounding arcs. This conversion corresponds to substituting out all latent variables from the structural equations of (3.3) and then constructing a new diagram by connecting any two variables $X_i$ and $X_j$ by (i) an arrow from $X_j$ to $X_i$ whenever $X_j$ appears in the equation for $X_i$ and (ii) a confounding arc whenever the same $\epsilon$ term appears in both $f_i$ and $f_j$. The result is a diagram in which all unmeasured variables are exogenous and mutually independent.

Several features should be noted from examining the diagrams in Figure 3.8.

[6] This was brought to my attention by James Robins, who has worked out many of these computations in the context of sequential treatment management (Robins 1986, p. 1423).

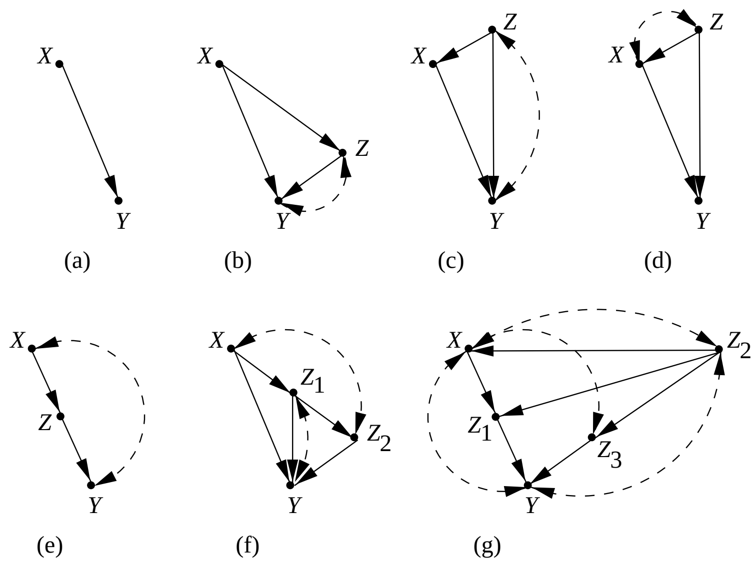


Figure 3.8: Typical models in which the effect of $X$ on $Y$ is identifiable. Dashed arcs represent confounding paths, and $Z$ represents observed covariates.

1. Since the removal of any arc or arrow from a causal diagram can only assist the identifiability of causal effects, $P(y|\hat{x})$ will still be identified in any edge subgraph of the diagrams shown in Figure 3.8. Likewise, the introduction of mediating observed variables onto any edge in a causal graph can assist, but never impede, the identifiability of any causal effect. Therefore, $P(y|\hat{x})$ will still be identified from any graph obtained by adding mediating nodes to the diagrams shown in Figure 3.8.

2. The diagrams in Figure 3.8 are maximal in the sense that the introduction of any additional arc or arrow onto an existing pair of nodes would render $P(y|\hat{x})$ no longer identifiable.

3. Although most of the diagrams in Figure 3.8 contain bow patterns, none of these patterns emanates from $X$ (as is the case in Figures 3.9(a) and (b) to follow). In general, a necessary condition for the identifiability of $P(y|\hat{x})$ is the absence of a confounding arc between $X$ and any child of $X$ that is an ancestor of $Y$.

4. Diagrams (a) and (b) in Figure 3.8 contain no back-door paths between $X$ and $Y$ and thus represent experimental designs in which there is no confounding bias between the treatment ($X$) and the response ($Y$); hence, $P(y|\hat{x}) = P(y|x)$. Likewise, diagrams (c) and (d) in Figure 3.8 represent designs in which observed covariates $Z$ block every back-door path between $X$ and $Y$ (i.e., $X$ is "conditionally ignorable" given $Z$, in the language of Rosenbaum and Rubin 1983); hence, $P(y|\hat{x})$ is obtained by standard adjustment for $Z$ (as in (3.21)):

$$P(y|\hat{x}) = \sum_z P(y|x, z)P(z).$$

5. For each of the diagrams in Figure 3.8, we readily obtain a formula for $P(y|\hat{x})$ by using symbolic derivations patterned after those in Section 3.4.3. The derivation is often guided by the graph topology. For example, diagram (f) in Figure 3.8 dictates the following derivation. Writing

$$P(y|\hat{x}) = \sum_{z_1,z_2} P(y|z_1, z_2, \hat{x})P(z_1, z_2|\hat{x}),$$

we see that the subgraph containing $\{X, Z_1, Z_2\}$ is identical in structure to that of diagram (e), with $(Z_1, Z_2)$ replacing $(Z, Y)$, respectively. Thus, $P(z_1, z_2|\hat{x})$ can be obtained from (3.45). Likewise, the term $P(y|z_1, z_2, \hat{x})$ can be reduced to $P(y|z_1, z_2, x)$ by Rule 2, since $(Y \perp\!\!\!\perp X|Z_1, Z_2)_{G_{\underline{X}}}$. We therefore have

$$P(y|\hat{x}) = \sum_{z_1,z_2} P(y|z_1, z_2, x)P(z_1|x) \sum_{x'} P(z_2|z_1, x')P(x'). \quad (3.50)$$

Applying a similar derivation to diagram (g) of Figure 3.8 yields

$$P(y|\hat{x}) = \sum_{z_1} \sum_{z_2} \sum_{x'} P(y|z_1,\ z_2,\ x')P(x'|z_2)P(z_1|z_2,\ x)P(z_2) \quad (3.51)$$

Note that the variable $Z_3$ does not appear in (3.50), expression above, which means that $Z_3$ need not be measured if all one wants to learn is the causal effect of $X$ on $Y$.

6. In diagrams (e), (f), and (g) of Figure 3.8, the identifiability of $P(y|\hat{x})$ is rendered feasible through observed covariates $Z$ that are affected by the treatment $X$ (since members of $Z$ are descendants of $X$). This stands contrary to the warning—repeated in most of the literature on statistical experimentation—to refrain from adjusting for concomitant observations that are affected by the treatment (Cox 1958; Rosenbaum 1984; Pratt and Schlaifer 1988; Wainer 1989). It is commonly believed that a concomitant $Z$ that is affected by the treatment must be excluded from the analysis of the total effect of the treatment (Pratt and Schlaifer 1988). The reason given for the exclusion is that the calculation of total effects amounts to integrating out $Z$, which is functionally equivalent to omitting $Z$ to begin with. Diagrams (e), (f), and (g) show cases where the total effects of $X$ are indeed the target of investigation and, even so, the measurement of concomitants that are affected by $X$ (e.g., $Z$ or $Z_1$) is still necessary. However, the adjustment needed for such concomitants is nonstandard, involving two or more stages of the standard adjustment of (3.21) (see (3.30), (3.50), and (3.51)).

7. In diagrams (b), (c), and (f) of Figure 3.8, $Y$ has a parent whose effect on $Y$ is not identifiable; even so, the effect of $X$ on $Y$ is identifiable. This demonstrates that local identifiability is not a necessary condition for global identifiability. In other words, to identify the effect of $X$ on $Y$ we need not insist on identifying each and every link along the paths from $X$ to $Y$.

### 3.5.2 Nonidentifying Models

Figure 3.9 presents typical diagrams in which the total effect of $X$ on $Y$, $P(y|\hat{x})$, is not identifiable. Noteworthy features of these diagrams are as follows.

1. All graphs in Figure 3.9 contain unblockable back-door paths between $X$ and $Y$, that is, paths ending with arrows pointing to $X$ that cannot be blocked by observed nondescendants of $X$. The presence of such a path in a graph is, indeed, a necessary test for nonidentifiability (see Theorem 3.3.2). That is not a sufficient

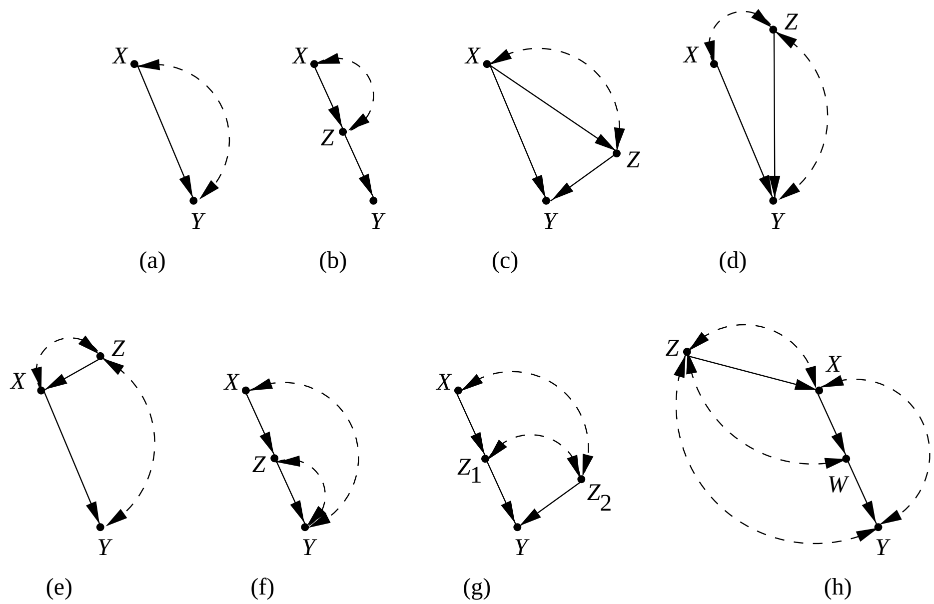


Figure 3.9: Typical models in which $P(y|\hat{x})$ is not identifiable.

test is demonstrated by Figure 3.8(e), in which the back-door path (dashed) is unblockable and yet $P(y|\hat{x})$ is identifiable.

2. A sufficient condition for the nonidentifiability of $P(y|\hat{x})$ is the existence of a confounding path between $X$ and any of its children on a path from $X$ to $Y$, as shown in Figures 3.9(b) and (c). A stronger sufficient condition is that the graph contain any of the patterns shown in Figure 3.9 as an edge subgraph.

3. Graph (g) in Figure 3.9 (same as Figure 3.7(c)) demonstrates that local identifiability is not sufficient for global identifiability. For example, we can identify $P(z_1|\hat{x})$, $P(z_2|\hat{x})$, $P(y|\hat{z}_1)$, and $P(y|\hat{z}_2)$ but not $P(y|\hat{x})$. This is one of the main differences between nonparametric and linear models; in the latter, all causal effects can be determined from the structural coefficients and each coefficient represents the causal effect of one variable on its immediate successor.