

Science Table Plots

Visualizing potential outcomes with Science Table plots

Randomization-assignment-control plots begin to suggest how the treatment assignment mechanism might be broken down into functionally different components. Here, we consider one way of breaking down the outcomes into two components: the control potential outcome, and the treated potential outcome. Doing so requires a shift in perspective. Until now, we’ve discussed the potential outcomes in terms of the prognostic score, which summarizes *baseline variation* important to the control potential outcome. This section will focus on the potential outcomes *themselves*, and for the moment we will shift the focus away from baseline variation.

The complete set of treated and control potential outcomes for all study individuals is sometimes referred to as the “Science Table” [cite Rubin]. Figure 1 contains a schematic of a visual representation of this concept, which we call a “science table plot.” Fundamental to the problem of causal inference is that the complete science table is unobservable, so visualizing the science table directly is only possible in theoretical and educational settings. However, the science table plot offers an interesting graphical view of the treatment effect at an individual and sample-wide level.

In a science table plot, individuals are depicted according to their treatment and control potential outcomes. In this setting, a subject’s individual treatment effect is graphically represented as their vertical distance to the diagonal line representing $Y(0) = Y(1)$ (Figure 1A), and different areas of the plot represent positive, negative, or neutral responses to treatment. Average treatment effect, then is the average distance of all points to the diagonal. Figure 1A-B. Visualizes two commonly assumed scenarios for treatment response: a relatively homogenous additive treatment effect, and a relatively homogenous negative one. Science Table plots offer an interesting visualization of the two major “null” hypotheses of causal inference. Figure 1C illustrates Neyman’s null hypothesis: the average treatment effect is zero, but individual treatment effects may still vary. This is in stark contrast to Fisher’s sharp null (Figure 1D), in treated and control potential outcomes are identical for all individuals in the study. This juxtaposition in particular may be an intuitive illustration in educational settings.

A science table plot can illustrate a variety of hypothetical models for treatment response, graphically depicting different forms of treatment effect heterogeneity. A few examples are shown in Figure 2. Multiplicative effects (Figure 2A) and tobit effects (Figure 2E) are two other treatment effect models occasionally discussed in the literature. Many discussions of treatment effect heterogeneity center on the hypothesis that treatment response correlates in some way with the control potential outcome, $Y(0)$: Perhaps “healthy” individuals with higher $Y(0)$ have a greater treatment response (Figure 2B), or perhaps “sicker” individuals with lower $Y(0)$ stand to gain more from treatment (Figure 2C). In other scenarios, one might suggest that there is a “sweet spot” for treatment, perhaps representing individuals who have intermediate potential outcomes (Figure 2D).

Figure 2F-H suggest other scenarios which may be more subtle and difficult to identify using common statistical approaches. Figure 2F depicts a scenario in which many individuals do not substantially benefit from treatment, but there are occasional large responses to treatment. This possibility is discussed further by Rosenbaum [cite DoOS] in his consideration of the National Supported Work Experiment, the subject of Lalonde’s 1986 study [cite lalonde]. Figure 2G suggests a different possible pattern for heterogeneity. Treatment response is highly heterogeneous, but not in a way that is associated with $Y(0)$. While treatment effect is zero on average, many people stand to benefit from treatment, and many people may respond quite badly. It’s possible that this heterogeneity can be explained in part by some measurable covariate independent of $Y(0)$. If so, it may be possible to identify subjects who stand to benefit from treatment and

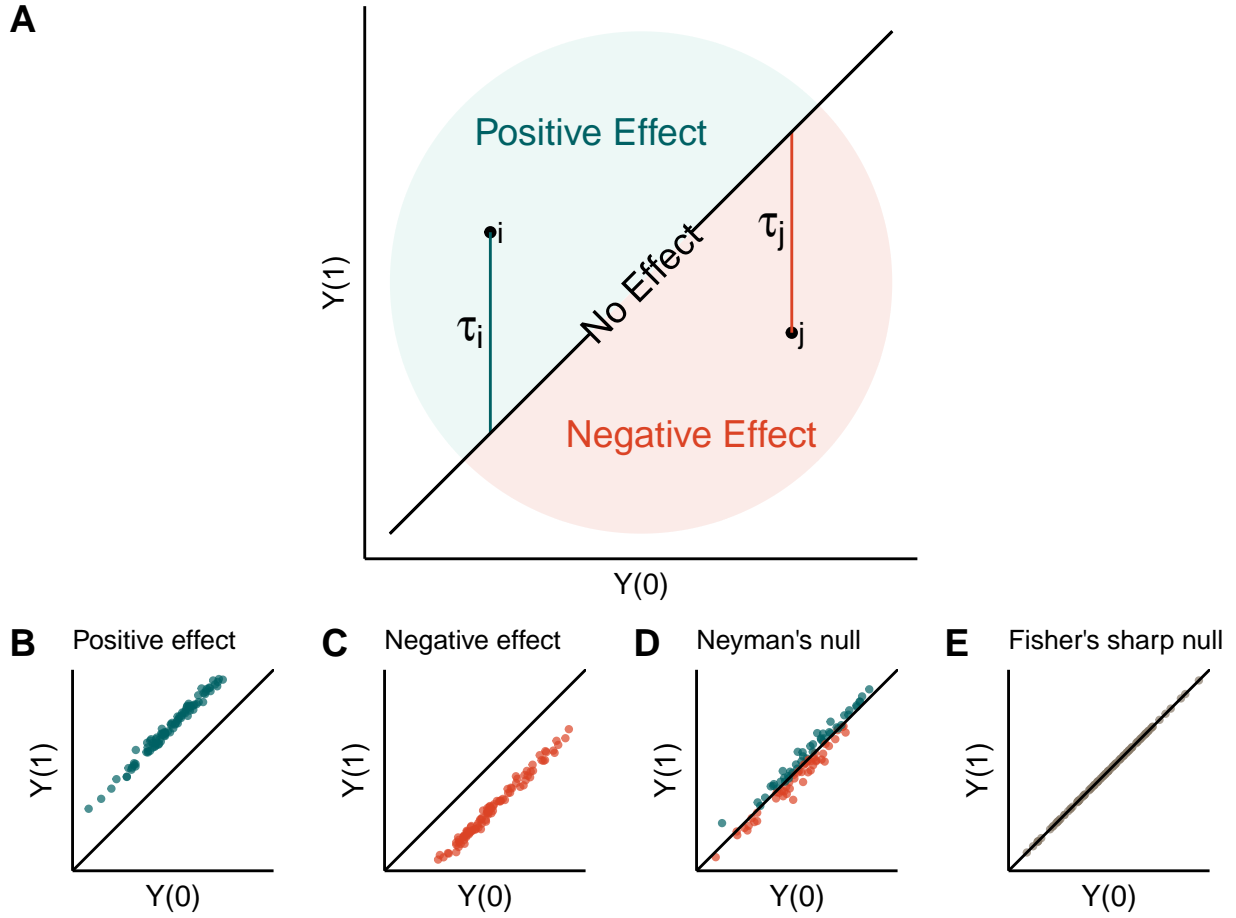


Figure 1: Science table plots. When a subject is graphically plotted according to their potential outcomes, the vertical distance to the diagonal represents their individual treatment effect, and different plot regions correspond to a positive, negative, or neutral response to treatment (A). Panels (A) and (B) depict a relatively homogenous positive or negative additive treatment effect, respectively. In Neyman's null hypothesis, treatment effect is zero on average, but there may still be some heterogeneity in individual responses (C). Fisher's sharp null (D) supposes that all individual level treatment effects are zero: graphically, this means all individual points on a science table plot lie on the diagonal.

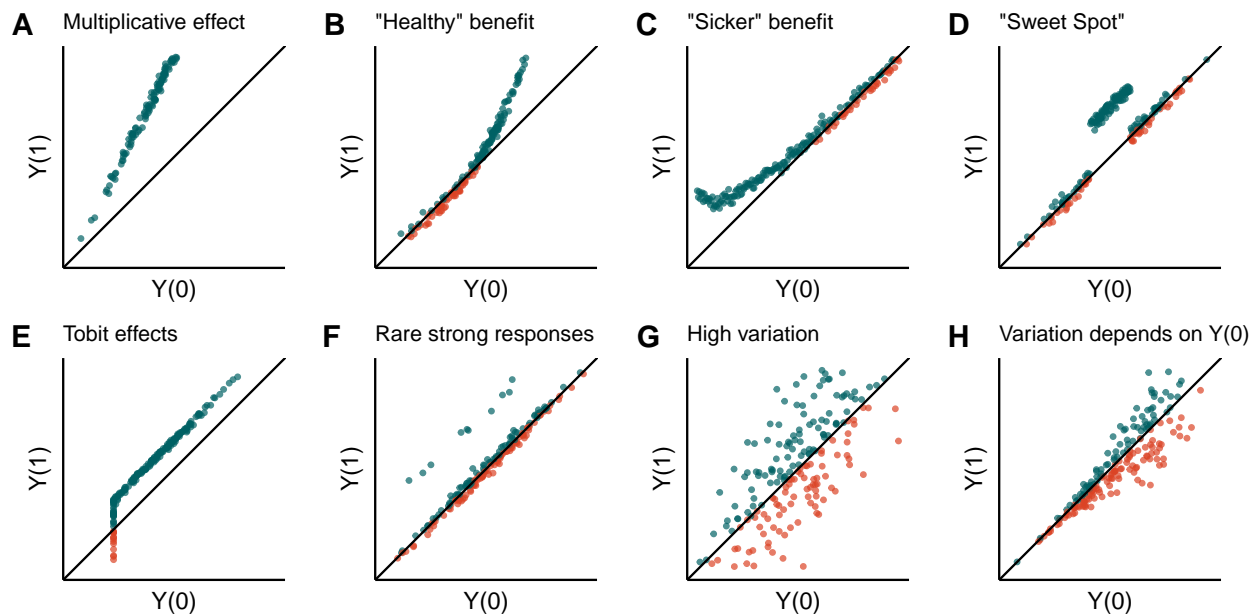


Figure 2: Science table plots depicting a variety of treatment response scenarios

those who will not. Otherwise, treatment response may be highly unpredictable, posing a difficult problem for implementation. Figure 2H considers a final possibility. In this scenario, the average treatment effect is zero at all levels of $Y(0)$, but treatment response heterogeneity itself is correlated with the control potential outcomes. Individuals with some $Y(0)$ have a very uniform response to treatment, while individuals with other levels of $Y(0)$ have highly variable responses. These examples suggest some ways that treatment effect heterogeneity may take unexpected forms which may be difficult to identify.