

# Science Table Plots Brainstorm

Rocky Aikens

2022-09-08

## 1 The Science Table Plot

Many a graduate-level course on causal inference begins with a discussion of the *Science Table*: the complete set of potential outcomes for all individuals in a sample (**rubin?**). The science table – and the potential outcomes within it – become the teaching foundation for a variety of causal inference approaches. Herein, I propose a new teaching visualization, *the science table plot*, which depict individuals according to their treated and untreated potential outcomes. While in practice the science table plot, like the science table itself, is fundamentally unobservable, in the classroom the science table plot is a handy companion to the science table, underscoring and clarifying foundational causal inference concepts.

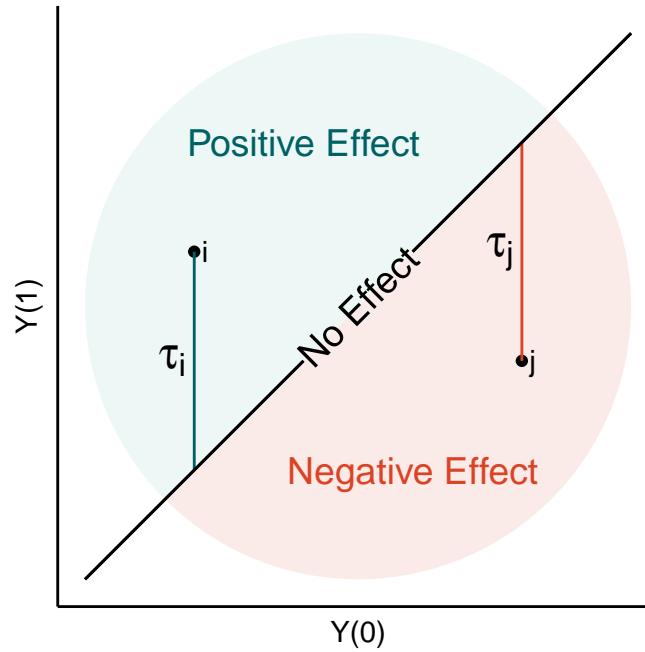


Figure 1: Science table plots. When a subject is graphically plotted according to their potential outcomes, the vertical distance to the diagonal represents their individual treatment effect, and different plot regions correspond to a positive, negative, or neutral response to treatment

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. Other estimands which focus on just a subset of the sample, such as the average

treatment effect among the treated, are simply the average distance to the diagonal for all points in the subsample.

The next plot visualizes two commonly assumed scenarios for treatment response: a relatively homogenous additive treatment effect, and a relatively homogenous negative one.

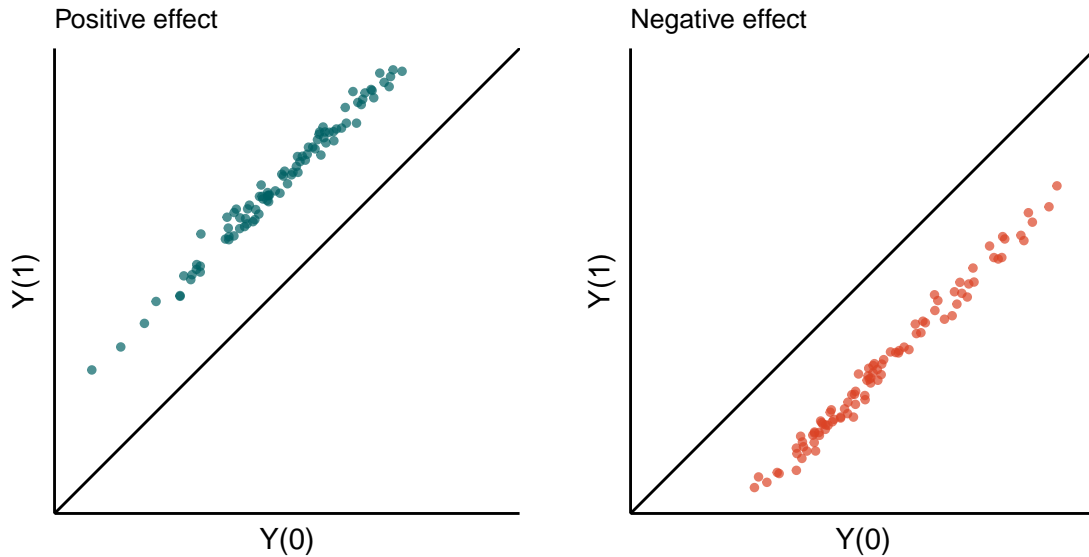


Figure 2: Panels (A) and (B) depict a relatively homogenous positive or negative additive treatment effect, respectively.

## 2 The null hypotheses

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.

## 3 The fundamental problem of causal inference

The fundamental problem of causal inference of course is that we can’t observe both potential outcomes at the same time. On a science table plot, this means that either we observe a subject’s vertical position (treated outcome) or their horizontal position. We might imagine that each point, then, becomes a line.

If treatment were entirely independent of the potential outcomes (as in a randomized trial), we could get a good approximation of the average treatment effect by calculating the average observed outcomes among the treated and control groups. Here, we see a Science Table plot for a sample where Fisher’s null is true: all points lie on the diagonal because each unit’s treated and untreated potential outcome are identical. This plot depicts an random 1:1 allocation scheme sorting units so that either their treated (horizontal line) or untreated (vertical line) outcomes are observed, as in a randomized trial. For each treated unit, only the treated outcome is observed, so rather than knowing the precise coordinates of their potential outcomes, we know only their horizontal position; they could be anywhere along the horizontal line defined by their treated potential outcome (and likewise, for controls). A standard difference-in-means approach takes the average

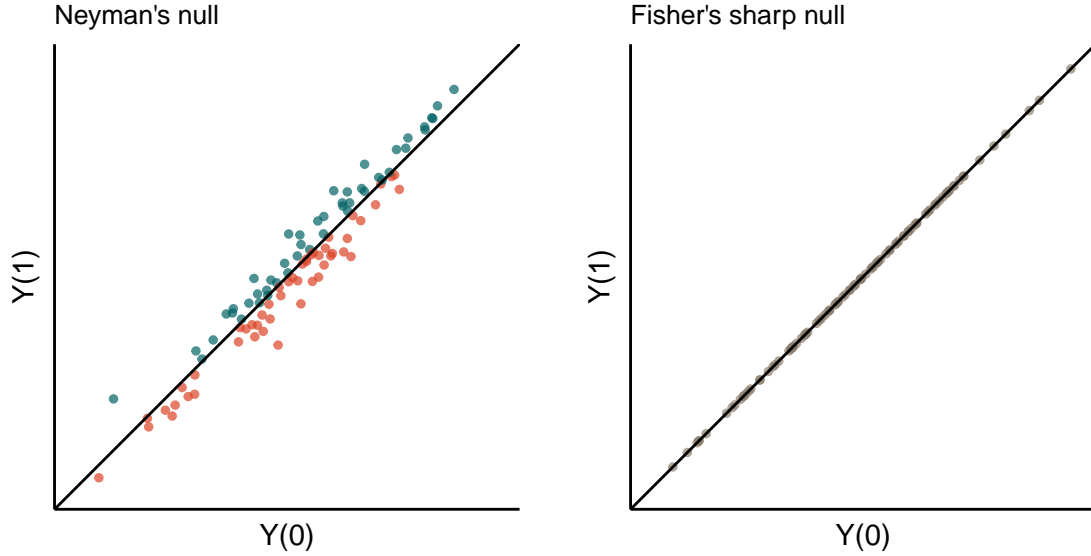
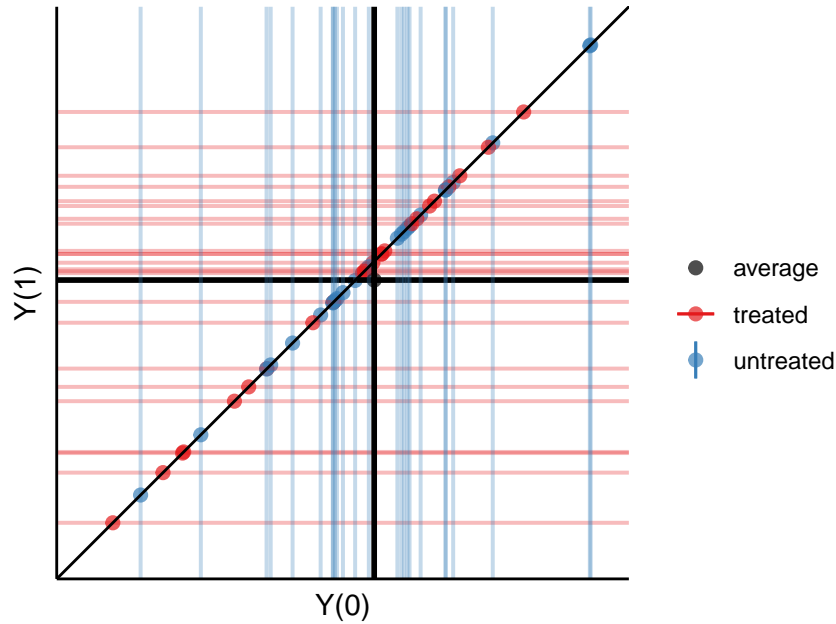


Figure 3: In Neyman's null hypothesis, treatment effect is zero on average, but there may still be some heterogeneity in individual responses (C). Fisher's sharp null (D) supposes that all individual level treatment effects are zero: graphically, this means all individual points on a science table plot lie on the diagonal.

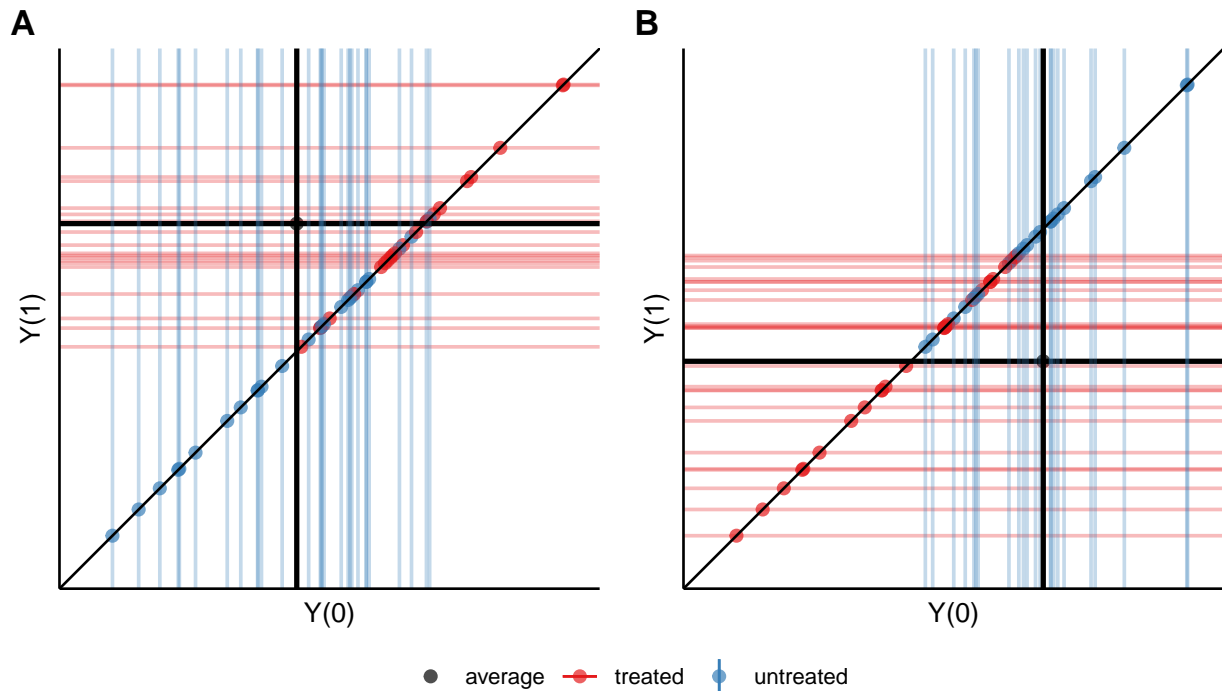
observed treated outcome (horizontal black line) and subtracts the average observed untreated outcome. Visually, the distance from the intersection of the black lines to the diagonal represents the difference in means estimator. Any deviation of the estimator from the true mean in this case reflects random variation resulting from the specific estimator, but not any bias due to non-random assignment.



In an observational study, things are not so simple. Here, we see that the same underlying science table (with the Fisher's exact null true), could produce either a positive or a negative result depending on who is allocated to treatment and who is not. This is selection bias.

```
## [1] 0.42
```

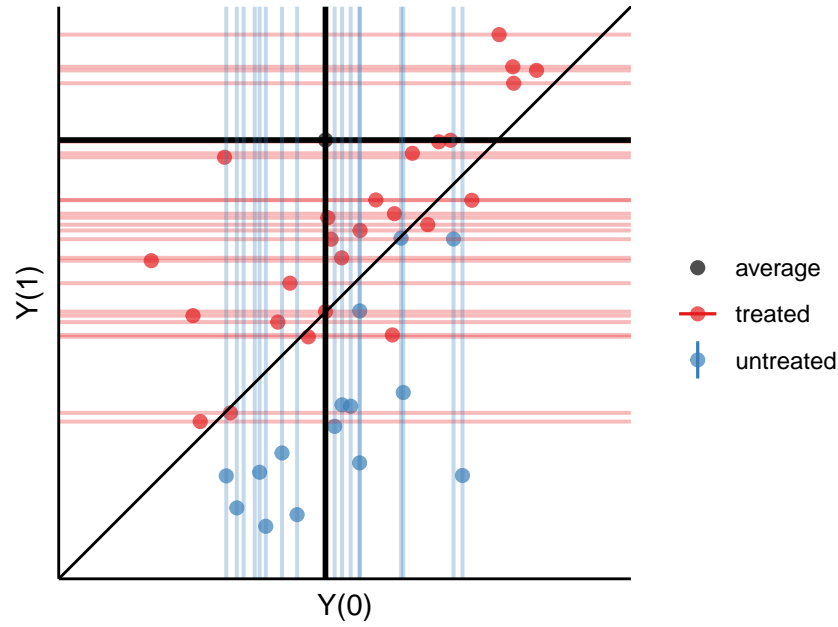
```
## [1] 0.52
```



### 3.1 Targeted selection

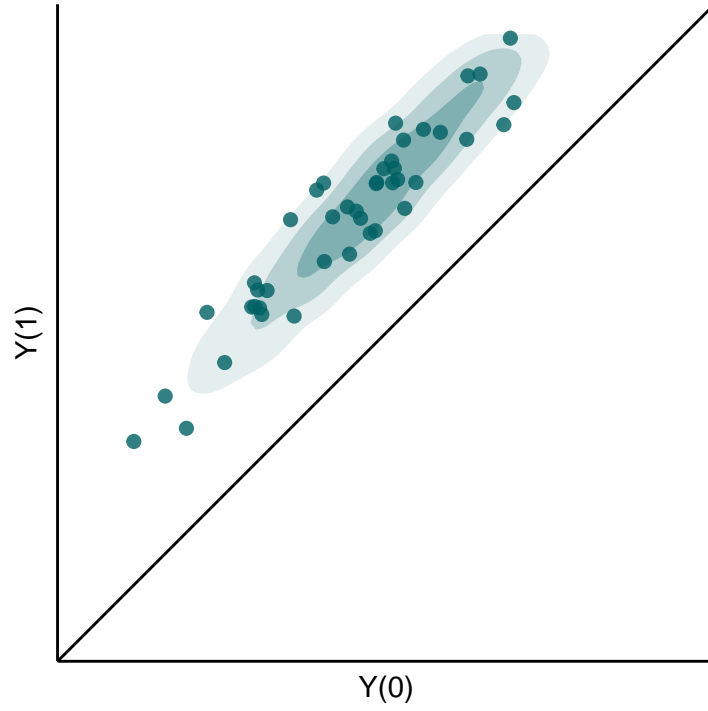
One scenario of particular concern in economics is the possibility of *targeted selection* the probability that an individual is in the treated group depends on their underlying treatment effect. The setting below is another null Science Table (this time a Neyman's null), in which there is a large heterogeneity in treatment effect and strong targeted selection. Individuals with more positive responses to treatment (above the diagonal) are more likely to select into the treated group. This can also cause bias, even if the treated and untreated units have the same untreated potential outcomes on average (as is the case here).

```
## # A tibble: 2 x 3
##   t      y0    y1
##   <lg1> <dbl> <dbl>
## 1 FALSE  2.31  1.03
## 2 TRUE   3.01  3.97
```



## 4 Superpopulation theory

In randomization inference, the sample is all there is. The sample average treatment effect is the average distance of all points to the diagonal. In superpopulation theory, the sample is itself made up of realizations from a probability distribution which determines the potential outcomes. If we imagined we could plot the density of the probability distribution over  $Y(0)$  and  $Y(1)$ , the average treatment effect would be (maybe?) the distance from the mean of this distribution to the diagonal. The sample is merely a window into the true distribution of interest.



## 5 Treatment effect heterogeneity

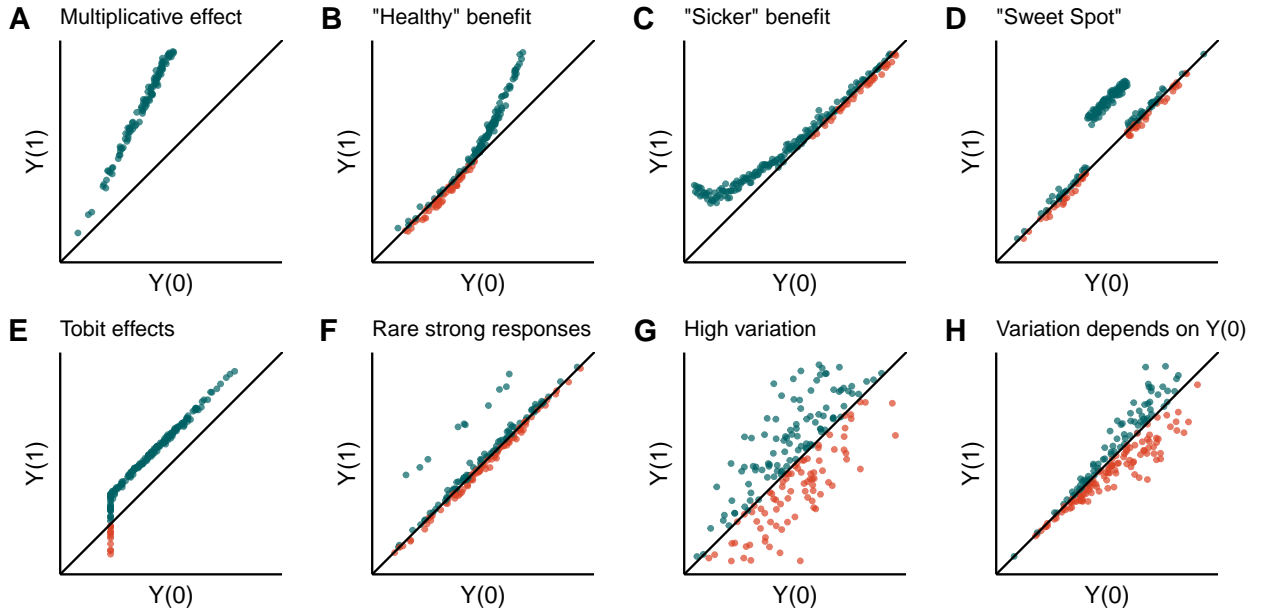


Figure 4: Science table plots depicting a variety of treatment response scenarios

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 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.