Chapter 16
Uncommon but Dramatic Responses to Treatment

Abstract Large effects in moderate to large studies are typically insensitive to small and moderate unobserved biases, but the concept of a ‘large effect’ is vague. What if most subjects are not much affected by treatment, but a small fraction, perhaps 10% or 20% of subjects, are strongly affected? On average, such an effect may be small, but not at all small for the affected fraction. Is such an effect insensitive to small and moderate unobserved biases?

16.1 Large Effects, Now and Then

Are large but rare effects insensitive to unmeasured biases?

In §2.5, in the National Supported Work Experiment, depicted in Figure 2.1, many men appear to have received little or no benefit from the treatment, but the few men with high earnings were fairly consistently in the treated group. In §2.5, if pairs of men were examined two at a time (i.e., four men), then 61% of the time, the man with the highest aligned earnings was a treated man, where 50% is expected by chance, but if pairs were examined 20 at a time (i.e., 40 men), 86% of the time the man with the highest aligned earnings was a treated man, where again 50% is expected by chance. Big gains in earnings are consistently in the treated group.

A similar pattern is seen for toxicity in Chapter 7 in Figure 7.1 in connection with Jeffrey Silber et al.’s [14] study of intensity of chemotherapy in $I = 344$ pairs of women with ovarian cancer: the median level of toxicity is not very different in the MO and GO groups, but high levels are more common in the MO group. If pairs of women are examined two at a time (i.e., 4 women), then in 65% of pairs, the woman with the highest aligned toxicity was treated by an MO, but if we look at 20 pairs at a time (i.e., 40 women), then 90% of the time, the woman with the highest aligned toxicity was treated by an MO. Figure 16.1 depicts the distribution of the MO-minus-GO differences, $Y_i$, in toxicity. In Figure 7.1 and Figure 16.1, it appears...
that many MOs and GOs treat with similar intensity producing similar toxicity, but
a fraction of MOs treat more intensively producing more toxicity.

In these cases, the hypothesis of a constant effect, $\tau = r_{Tij} - r_{Cij}$ for all $i, j$, does not look plausible. As discussed in \S2.5, a more plausible hypothesis is that $r_{Tij} - r_{Cij}$ is zero or small for many subjects $i, j$, but $r_{Tij} - r_{Cij}$ is large for some subjects $i, j$. David Salsburg [11] argued that effects of this sort are fairly common, are often important, and that we tend to miss them because our methodology tends to focus on typical effects, but in this context large effects do occur but are not typical. The question addressed in the current chapter is whether an effect of this kind is highly sensitive to unmeasured biases. In both Figure 2.1 and Figure 7.1, the typical difference in outcomes is not large, yet the distributions of outcomes are quite different.