

What Longo and Drazen should have written

Carl T. Bergstrom

At present the allocation of research effort within science seems to exhibit a reasonable balance between *data production* and *data analysis*. Traditionally, researchers who generate primary data have been permitted to hoard those data for extended periods of time, to the private advantage of their own research groups. Today, with the frequent calls for data sharing and occasional policies for data deposition, researchers are finding it more difficult to hoard data.

The consequence of this trend is a reduction in the cost of obtaining others' data, coupled with a reduction in the benefits of producing one's own. One doesn't need a PhD in economics to see what the outcome will be: researchers will decrease effort allocated toward generating data, and increase effort toward analyzing it.

We don't want to allow data hoarding, so we're stuck with the decreasing cost of data acquisition. If we want to maintain the previous balance of the two activities, we need to find some way of increasing the rewards associated with data creation. More formal mechanisms of acknowledging and rewarding data use—perhaps something intermediate between citation and authorship—would be a good start.