J. Jason Bell

June 5, 2013

I recently read two books. First I read *Uncontrolled*, by Jim Manzi, and then I read *The Signal and the Noise* by Nate Silver. Perhaps the key point of *Uncontrolled* is that, in the quest for "useful, non-obvious, predictive rules" experiments are king, particularly iterative ones. *The Signal and the Noise* on the other hand, probably focuses most on arguing for the value of a Bayesian way of thinking when making forecasts. Rather than giving two separate synopses, I want to lay out the books side by side for comparison.

Both books are concerned with prediction, in some sense. The writing suggests that both authors gained their conviction from experience rather than from education. Specifically, both of them were practitioners rather than academics. Both are very familiar with academic findings and style, however. Manzi's conclusion is that useful predictive rules can be discovered via experiments, but by basically nothing else. He criticizes pattern-finding algorithms, regression, and pure theory. His proposed solution to almost everything is the experiment. This is, after all, the gold standard in the academic world for establishing causality. Silver is mostly concerned with forecasting. His conclusion is that forecasters should think in a Bayesian way, by which he means we should establish prior beliefs about the probability of an event occurring and then we should use Bayes' rule to update those priors when we gain new information. Silver also argues that humans should incorporate theory into forecasting, albeit in a very cautious way. Silver doesn't consider experiments at all, though this is mostly because experiments are mostly impossible in the applications he discusses: earthquakes, disease, the stock market, weather, poker, baseball, political campaigns, etc.

It may be obvious now that, while both are concerned with prediction, the types of prediction are fairly different. Manzi is concerned with predicting the consequences of human intervention in a process. His main applications are marketing: if I own a store and I increase the price of milk, how many people will go elsewhere? Silver wants to forecast processes which are too big for one person to manipulate: who will be elected president. The two tasks are extremely different. Manzi argues pretty convincingly that Silver's techniques don't add any value to experiments when applied to his applications, and Silver's applications are impossible to control, and experiments are not applicable.

The two books do occasionally touch the same topics, however, and when they do discrepancies tend to arise. For example, Manzi investigates which crime fighting techniques seem to work the best, and he concludes that, while almost nothing works well, broken-

windows policing, where officers target the environment rather than major crimes, is the most promising. Silver, on the other hand, says "the empirical evidence for the merit of this theory is quite mixed." These aren't wildly different conclusions, after all, Manzi doesn't claim that broken-windows works for sure, but merely says that it is the only policing technique that stands up to experimental scrutiny to any degree, and Silver simply says the evidence is mixed. Still, the difference exists. The tone of Manzi is that broken-windows is promising, while Silver suggests that broken-windows is embraced because it's easy for police officers to do.

Both books discuss the 2008 financial crisis, of course. Manzi uses the experience to illustrate the futility of statistical modeling and the depth of human overconfidence and bias, while Silver points out the number of people who called the crisis correctly.

One discrepancy which seems to exist between Manzi, Silver, and a man named Robin Dawes, is the value of human judgement. Dawes conducted many experiments which pitted human judgements against regression analysis. For example, doctors saw patients and were asked whether the patients had certain conditions. At the same time, each patient's information was used in a linear regression to predict the chance of the same conditions. The regression always predicted more accurately, when averaging over many patients. Dawes performed these experiments in many different contexts, from medicine to job interviews to stock picking. Yet, both Manzi and Silver defend the value of human judgements when making forecasts. Manzi goes so far as to say that experienced humans are better at predicting things like the market-share of a new restaurant than a regression would be. Silver says that the use of computer models (including regression) can be significantly improved by incorporating theoretical foundations for the variables in question.

I think the discrepancies above can be reconciled, at least to some degree. My reconciliation takes the form of the following two-by-two table:

|                  | Controllable          | Uncontrollable    |
| ---------------- | --------------------- | ----------------- |
| Stable, Simple   | Data Only             | Data Only         |
| Complex or Chaotic | Experiments and Theory | Data and Theory  |

The difference between controllable and uncontrollable is self-evident. Some examples: you can't create hurricanes, but you can change a business logo. As other axis, this requires a bit more explanation. What is meant by a stable, simple process (SS from now on) is that the underlying process is the same from case to case. Diagnosing a disease which has common indicators is better done by a computer because it won't get tired, won't take shortcuts, and won't use heuristics. It will treat every case exactly the same, which is the best thing to do for a SS process. Complex or Chaotic processes cannot be defined easily by a few indicators. Chaotic processes are time dependent and non-linear. The effects often change at exponential rates. Hurricanes are an example of these. Complex processes

are sometimes SS, but they can also do erratic things. You might be tempted to use SS methods on them, and then call the erratic events errors. The problem here is that the erratic events are sometimes extremely important. Earthquakes are thought to be complex processes, and the erratic events are 9.0 earthquakes, so you have to care about them.

What is meant by the term "data only" in the top cells is basically that the data can mostly speak for itself, and simple statistical models can be applied to these problems. Linear regression is a very simple statistical model and can be used in a large variety of circumstances, and can replace human judgement for many mundane tasks.

The bottom left cell is meant to suggest that when you can control the variable well, experiments are best. Experiments and theory are almost inseparable, so the use of both is natural.

The bottom right cell is meant to suggest that, when processes are complex or chaotic, human judgement is likely to add more value. Humans are creative, while computers are not. Simple statistical models fail to predict complex or chaotic processes. Eventually, as humans learn more about these processes, some of them can be moved to the SS squares. Medicine has many examples of this.

This table is coarse, and so it is important to clarify one thing: in the bottom two cells, the incorporation of theory is only optimal when the value added by human thought is less than the bias added. This is obvious in principle but not in practice. How can you tell when the value is greater than the bias? Both Manzi and Silver suggest that you iterate. In addition, the difference between controllable and uncontrollable is not always obvious. Often, you can control one piece of the system. In these cases, it seems that heavy use of data-mining (computer algorithms that look for patterns in the data) are good for short term prediction, but bad for understanding predictive rules. Theory is better for longer term understanding of the phenomenon, which eventually leads to better prediction, but generally theories don't do nearly as well in the short term as algorithms do.

I have to disclaim the fact that my thoughts on this subject are still in flux. You may come to different conclusions about these things. I'd love to hear your thoughts.