The Seven Properties of Good Models[1]

Xavier Gabaix
MIT and NBER

and

David Laibson
Harvard University and NBER

NYU Methodology Conference

Scientists spend most of their time formulating and analyzing models. A model is a description – or representation – of the world. Most models are based on assumptions that are known to be only approximately true (and exactly false). For example, consider the most commonly used models of the earth: flat, spherical, and ellipsoid. These models do not account for the bumps and grooves. A perfect replica of the earth would reproduce every contour, but such a representation would be impractical. You don't need to know the height of Beacon Hill to take a subway across Boston. Tourists use a flat subway map -- the model that is just complex enough for the problem at hand.

This essay describes the seven key properties of useful economic models: parsimony, tractability, conceptual insightfulness, generalizability, falsifiability, empirical consistency, and predictive precision.[2] Successful economic models have most of these properties, although almost no economic models have them all. Some of these seven properties are already well accepted among economists, specifically, parsimony, tractability, conceptual insightfulness, and generalizability. The other properties -- falsifiability, empirical consistency, and predictive precision – are not universally accepted.

We believe that these seven properties are fundamental. Some economists instead argue that classical optimization assumptions -- like rationality and dynamic consistency -- are necessary ingredients of a good economic model. We believe that these optimization assumptions do discipline economic analysis and often *produce* many of the

---

[2] Many other authors have attempted to characterize the purposes of models. Our formulation has been directly and indirectly influenced by preceding analyses by Popper [1935, 1963], Merton [1949], Friedman [1953], Kuhn [1962], and Stigler [1965]. For a related contemporary analysis, see Jasso [2004]. Stigler [1965] is the most important precedent for the analysis in the current paper. Stigler identifies three characteristics of economic theories that are accepted by "leading economists." Stigler's criteria are generality, manageability, and congruence with reality.

seven properties listed above. For instance, parsimony, tractability, conceptual insightfulness, and generalizability all tend to follow from classical optimization assumptions. However, such classical assumptions are not the only path to a good model. We believe that classical optimization assumptions are better treated as hypotheses that should be tested and not fundamental or necessary properties of economic models. Anticipating one set of objections, we also discuss how to conduct normative analysis even if economists can't rely on optimization principles like revealed preferences.

We conclude the essay by discussing the appropriate empirical scope of economic models. We argue that restrictions on the scope of economic research contradict the history of science, as well as the conceptual orientation of economics itself.

Seven key properties of economic models.

Economic models are conceptual frameworks that aid in the understanding, description, and/or prediction of human behavior. Formal empirical analysis, casual observation, and introspection all play a role in demonstrating the correspondences between economic models and the world.

In this section, we define the seven key properties of good models. We note that these properties are not always mutually consistent. Consequently, these properties sometimes need to be traded off against each other.
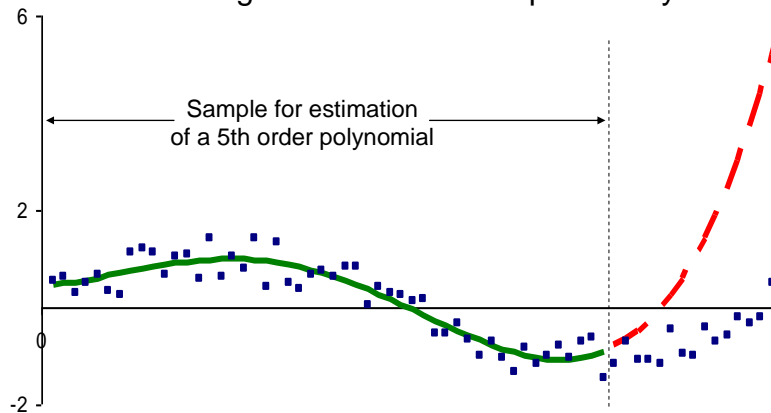
1. Parsimony
2. Tractability
3. Conceptual insightfulness
4. Generalizability
5. Falsifiability

6. Empirical consistency
7. Predictive precision

Parsimonious models are simple models in the sense that they rely on relatively few special assumptions and they leave the researcher with relatively few degrees of freedom.[3] Parsimonious models are desirable because they prevent the researcher from consciously or subconsciously manipulating the model so that it over-fits the available facts. Over-fitting occurs when a model works very well in a given situation, but fails to make accurate out-of-sample predictions. For example, if a model incorporates a *large* set of qualitative psychological biases then the model is nonparsimonious, since selective combination of those biases will enable the researcher to tweak the model so that it "explains" almost any pattern of observations. Likewise, if a model has many *free* parameters – for instance, a complex budget constraint or complex household preferences -- then the model is relatively nonparsimonious. When models are flexible and complex the researcher can combine the myriad elements to match almost any *given* set of facts. Such flexibility makes it easy to explain in-sample data, producing the *false* impression that the model will have real (out-of-sample) explanatory power (see Figure 1).

---

[3]These ideas are reflected in Occam's Razor, a principle attributed to the logician William of Ockham: "All things being equal, the simplest solution tends to be the best one."

Figure 1: The value of parsimony.

The data (squares) is generated by sin(x/10) + ε, where ε is distributed
uniformly between -½ and ½. The sold line fits the first 50 data points to
a fifth-order polynomial – a non-parsimonious model. The polynomial
has good fit in sample and poor fit out of sample (dashed line).

Tractable models are easy to analyze. Models with *maximal* tractability can be solved with analytic methods – i.e. paper and pencil calculations. At the other extreme, *minimally* tractable models cannot be solved even with a computer, since the necessary computations/simulations would take too long. For instance, optimization is typically not computationally feasible when there are dozens of continuous state variables – in such cases, numerical solution times are measured on the scale of years or centuries.

C*onceptually insightful* models reveal fundamental properties of economic behavior or economic systems. For example, the model of concave utility identifies the key property of risk aversion. The concept of concave utility is useful even though it makes only qualitative predictions. An optimizing framework like Nash Equilibrium is also conceptually insightful even though it relies on an assumption that is empirically false – perfect rationality. The concept of Nash Equilibrium clarifies some abstract ideas about equilibrium that are important to understand even if the Nash framework is an

incomplete explanation of real-world behavior. Finally, many models are conceptually useful because they provide normative insights.

Generalizable models can be applied to relatively wide range of situations. For example, a generalizable model of risk aversion could be used to analyze risk aversion in settings with small or large stakes, as well as risk aversion with respect to losses or gains. A generalizable model of learning could be used to study learning dynamics in settings with a discrete action set or a continuous action set or an action set with a mixture of discrete and continuous actions. A generalizable model of intertemporal choice could be used to study decisions with consequences that occur in minutes or decades.
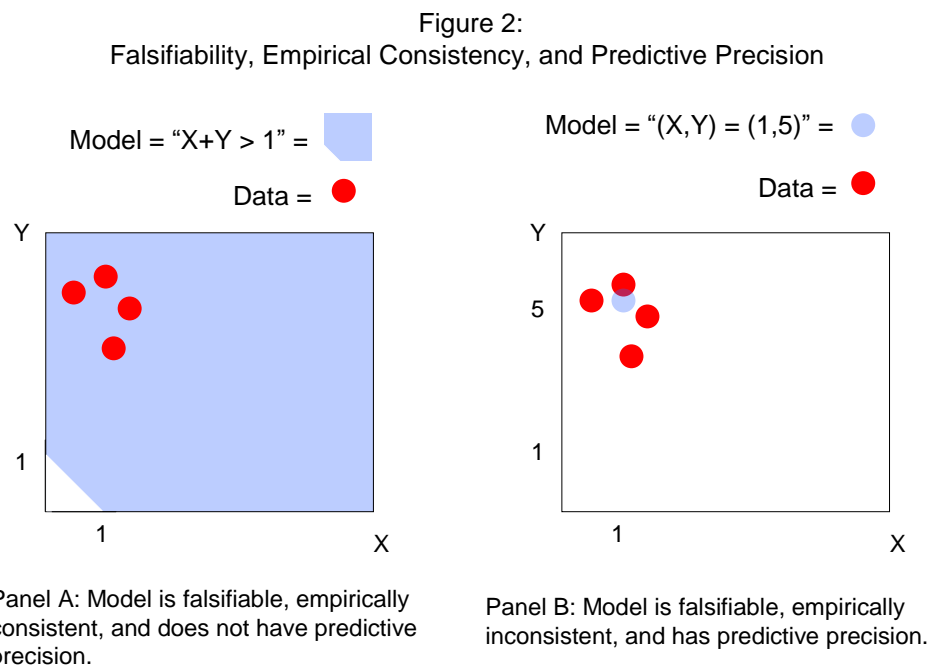
Falsifiability and prediction are the same concept. A model is falsifiable if and only if the model makes nontrivial predictions that can in principle be empirically falsified. If a model makes no falsifiable predictions, then the model can not be empirically evaluated.

Empirically consistent models are broadly consistent with the available data. In other words, empirically consistent models have not yet generated predictions that have been falsified by the data. Empirically consistent models can be ranked by the strength of their predictions. At one extreme, a model can be consistent with the data if the model makes only weak predictions that are verified empirically. At the other extreme, models can achieve empirical consistency by making many strong – i.e. precise -- predictions that are verified empirically.

Models have *predictive precision* when they make precise – or "strong" – predictions. Strong predictions are desirable because they facilitate model evaluation and model testing.   When an incorrect model makes strong predictions it is easy to

empirically falsify the model, even when the researcher only has access to a small amount of data. A model with predictive precision also has greater potential to be practically useful if it survives empirical testing. Models with predictive precision are useful tools for decision makers who are trying to forecast future events or the consequences of new policies.

A model with predictive precision may even be useful when it is empirically *inaccurate*. For instance, policymakers would value a structural model that predicts the timing of recessions, *even* if the model usually generated small predictive errors. An alternative model that *correctly* predicted that a recession would occur at some unspecified time over a ten year horizon would not be as useful. In general, models that make approximately accurate *strong* predictions are much more useful than models that make exactly accurate weak predictions. Figure 2 provides a visual illustration of this point.

Figure 2:
Falsifiability, Empirical Consistency, and Predictive Precision

Panel A: Model is falsifiable, empirically consistent, and does not have predictive precision.

Panel B: Model is falsifiable, empirically inconsistent, and has predictive precision.

Predictive precision is infrequently emphasized in economics research. Academic economists have instead elevated properties like parsimony, tractability, conceptual insightfulness, and generalizability. We believe that this tendency arises because people tend to celebrate the things they do best. Economists have had a comparative advantage in developing elegant mathematical models. Economists – including behavioral economists -- have been far less successful in developing general models that make precise *quantitative* predictions that are approximately empirically accurate. In this sense, economic research differs from research in the natural sciences, particularly physics. We hope that economists will close this gap. Models that make weak predictions (or no predictions) are limited in their ability to advance economic understanding of the world.

Is Optimization a Necessary Ingredient for an Economic Model?

We have summarized seven key characteristics of a good model. Some economists have argued that modeling criteria like those that we have discussed are incomplete. These economists formally *define* an economic model as a mathematical representation that has many of the features above *and* certain axiomatic optimization properties. For instance, rational beliefs and dynamic consistency have been proposed as axioms that define what it means to do economics. We do not agree with such a formulation.

We believe that economics is ultimately like any other science. Scientific models are mathematical approximations of the world. In every other scientific field, these

approximations are judged by something akin to the seven criteria listed above and not by a model's adoption of particular axiomatic assumptions.

In the history of science, every axiomatic litmus test has been discarded. In retrospect, it is easy to see that axioms like the flat earth, the earth at the center of the universe, or the Euclidean structure of space, should *not* have been viewed as inviolate. It is likely that every axiom that we currently use in economics will suffer the same fate that these earlier axioms suffered. Indeed, it is now believed that no branch of science will *ever* identify an inviolate set of axioms [Kuhn 1962]. This does not mean that we should abandon axioms, only that we should *not* use axiomatic litmus tests to define a field of scientific inquiry.

Relaxing economists' commitment to optimization axioms poses several problems. Optimization has provided discipline – i.e. parsimony. But optimization is neither necessary nor sufficient for parsimony. For example, physics models are not constrained by optimization and are nevertheless highly parsimonious. Moreover, there are many *optimization* models that are not parsimonious since they make many special assumptions – about budget constraints and preferences -- to explain a single behavioral regularity.

Without optimization axioms, economists will not able to rely on traditional normative tools like revealed preference. Instead, economists must develop models in which true preferences interact with other factors -- like biases, errors, and dynamic inconsistency -- to produce economic behavior. Economists will need to simultaneously model true preferences – hereafter called *normative preferences* -- and the confounding factors that prevent individuals from maximizing these normative preferences. Such

integrative models are not new. Since the work of Luce [1959], economists have been developing formal mathematical models that incorporate both normative preferences and decision-making errors.[4]

To impute normative preferences, economists should adopt a two-step strategy that generalizes the classical revealed preference framework. First, specify a model that includes both normative preferences and a positive theory of behavior (incorporating factors like decision-making errors and/or dynamic inconsistency which force a wedge between revealed preferences and normative preferences). Second, use choice data to estimate the model, thereby imputing the latent normative preferences.[5]

What is the Appropriate Scope of Economic Research?

All natural sciences have developed by incorporating the study of smaller and smaller units of analysis. Biologists began with organisms, and then studied cells, cellular organs, and molecules (most importantly DNA). Physicists and their intellectual precursors began with visible objects and subsequently studied atoms and sub-atomic particles. Similar intellectual progressions have also occurred in the social sciences. For instance, psychologists have recently embraced both neuroscience and molecular genetics. Some anthropologists are also using these new methods.

We believe that economic research will also incorporate smaller units of analysis. Economics is usually described as the study of the allocation of scarce resources.

---

[4] See Bernheim and Rangel [2006] – which is also published in this volume -- for an extended discussion of these issues.
[5] See Beshears et al [2006] for a more detailed discussion of these issues.

Economists identify the individual decision-maker as the key actor in this resource allocation process. Neuroscience[6] and cognitive genetics[7] will enable economists to develop a more complete understanding of the decision-making process. Ultimately, such research will advance economists' goal of predicting human behavior. Neuroscience and genetics data is not a *necessary* ingredient for economic research, but such biological data will speed up the economic research process by providing convergent types of evidence that complement traditional behavioral studies. These new data sources will enable economists to more quickly formulate, develop, and test models of decision-making and inter-individual variation.

Conclusion

This essay describes the seven key properties of successful economic models: parsimony, tractability, conceptual insightfulness, generalizability, falsifiability, empirical consistency, and predictive precision. However, even highly successful models do *not* have all seven properties. Many of the properties are in conflict with one another. For example, generalizing a model sometimes makes a model unfalsifiable -- the most general form of the theory of revealed preference can't be rejected by behavioral data.

We encourage economists to give more weight to the property of predictive precision. Models that make quantitatively precise predictions are the norm in other

[6] See Camerer et al [2004] for a recent review.
[7] Ding et al [2006], Benjamin et al [2007].

sciences. Models with predictive precision are easy to empirically test and when such

models are approximately empirically accurate they are likely to be useful.

Bibliography

Benjamin, Daniel, J., Christopher F. Chabris, Edward Glaeser, Vilmundur Gudnason, Tamara B. Harris, David Laibson, Lenore Launer, and Shaun Purcell. 2007. "Genoeconomics." Forthcoming in [title to be announced, will enter it in galleys], ed. Maxine Weinstein. National Academy of Sciences.

Bernheim, B. Douglas and Antonio Rangel. 2006. "Title Here," Forthcoming in [title of current volume], ed. Andrew Caplin and Andrew Schotter.

Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian. 2007. "How Are Preferences Revealed?" mimeo, Harvard University.

Camerer, Colin, Jonathan Cohen, Ernst Fehr, Paul Glimcher, David Laibson, George Loewenstein, Read Montague. 2007. "Neuroeconomics." Forthcoming in *The Handbook of Experimental Economics,* ed. John Kagel and Alvin Roth.

Ding, Weili, Steven Lehrer, J. Niels Rosenquist, and Janet Audrain-McGovern. 2006. The Impact of Poor Health on Education: New Evidence Using Genetic Markers. *NBER* Working Paper 12304.

Friedman. 1953. *Essays in Positive Economics*. Chicago: University of Chicago Press

Jasso, Guillermina. 2004. The Tripartite Structure of Social Science Analysis. *Sociological Theory.* 22:3, 401-431.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Merton, Robert K. 1973. *The Sociology of Science*. Chicago: University of Chicago Press.

Popper, Karl R. 1934, republished in English 1959. *The Logic of Scientific Discovery*. New York: Basic Books.

———. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books.

Stigler, George J. 1965. *Essays in the History of Economics*. Chicago: The University of Chicago Press.