

Assessing and Managing Model Risk

By Tom Coyne

As I define it, a model is the simplified representation of a dynamic situation that helps us focus attention on key variables, understand causal processes, and predict future conditions and the potential results of our actions. This definition includes not only quantitative models of financial and physical stocks and flows, but also qualitative models that utilize expert judgment and project and program plans, which model sequences of input activities and decisions, and predict their expected results. On the basis of this definition, it is easy to see that models are critical to business success, and therefore an important source of business risks. Yet too many enterprise risk management programs fail to adequately assess and manage these risks. In this short overview, I have two goals: to provide you with a better understanding of the nature of these risks, and improve your ability to address them.

Sources of Model Risk

There are a number of sources of model risk that organizations need to address. At the simplest, mechanical level, input values and relationships may be incorrectly entered into a model. For example, a nine may be entered instead of a zero, formulas can be entered incorrectly, or a finish/start dependency overlooked in a project plan. The good news is that these are also the easiest mistakes to catch and correct.

In many cases, the values for model inputs must be estimated, which can give rise to estimation errors. For example, even when we are dealing with a normal (also known as Gaussian or bell-curve) distribution, the mean is estimated with a standard error (equal to the standard deviation divided by the square root of the

number of observations used to estimate the mean). In so far as we as ERM managers are concerned with exposures in the tail(s) of the distribution (i.e., in the range defined by the estimated mean plus or minus two or more standard deviations), these standard errors can have a very large, if often unrecognized, impact on our decisions. Similarly, when formulating a project plan/model, a manager must make estimates of the length of time and resources that will be required to complete different tasks. If this is done qualitatively, then the manager's estimation error can result from overoptimism, overconfidence, anchoring, and other common cognitive biases. Alternatively, if these estimates are derived from quantitative analysis, estimation error can result from using the wrong distribution to formulate a single point estimate for the input value, or to describe the shape of the range of possible outcomes if a Monte Carlo simulation model is being used. Mathematically, the Gaussian or normal (i.e., "bell curve") distribution is the most tractable. Unfortunately, research has shown that outcomes produced by complex adaptive systems of interacting human beings (such as commodity or financial markets, or project execution), as well as non-linear physical processes (like fires or weather) are not Gaussian. Rather, they tend to follow a power law distribution, and have "fatter" and "longer" tails than the normal bell curve.

In addition to errors related to the estimation of input parameter values, another critical source of model error is the possibility that the model itself has left out key variables, or misspecified the nature of one or more of the relationships between the variables it contains. In many real world situations, these relationships are very complex, and involve non-linearities and time-delays. Traditionally, many models have used the correlation statistic to describe the relationships between key variables. However, correlation suffers from critical limitations, as it only measures the linear relationship between the average levels of two variables. This can lead to substantial underestimates of the true level of risk, particularly in the case of extreme events that are usually of the most interest to risk managers.

The limitations of correlation have led to the use of more sophisticated quantitative techniques such as copulas. However, deciding which copula to use to describe the relationship between variables is often made very difficult by sparse or absent historical data about the occurrence of extreme events.

A more fundamental source of model error is the difficulty we have as human beings with causal reasoning. Most of what we are taught in school about causation comes from the physical sciences, where: (1) Observed effects usually have relatively few causes; (2) Cause and effect are usually related by unchanging natural laws; and (3) Experiments can be repeated to test and confirm hypotheses. In practice, however, many of the causal reasoning situations we encounter involve “complex adaptive systems”, where: (1) Effects often have multiple causes, and causal relationships are often non-linear (i.e., characterized by positive feedback loops) and time-delayed; (2) This produces effects that are “emergent” rather than the result of intentional causation; (3) Causal relationships also change over time, as the agents involved in the system adapt their behavior and relationships in light of the results produced by their actions; and (4) Experiments are difficult to conduct and replicate. Causal reasoning about complex adaptive systems will inevitably be imperfect. The best we can hope for is a “coarse grained” understanding that will have an irreducible level of uncertainty. For example, complex adaptive systems typically produce a power law (exponential) distribution of effects – i.e., many small ones, and few big ones. They are also often characterized by three different behavioral regimes (phases) – one is very stable, one is very unstable, and one in the middle maximizes system resilience and adaptability. This gives rise to a familiar pattern, with a large change usually preceded by a number of smaller events of lesser magnitude, that hint at the pressure for change that is building in the system. Rapid change often occurs once one or more system “control parameters” have passed a critical threshold, or “tipping point” (e.g., the amount of leverage in a financial system). This process creates the familiar “S-shaped” pattern of change in many domains.

Managing Model Risk

Broadly speaking, there are two steps involved in effectively managing model risk. The first is model verification. This set of activities focuses on whether a model has been properly documented, and whether its mechanical operation is sound (e.g., no errors in entries describing input parameter values and the relationships between different inputs and outputs). Most people who build models focus their quality control/risk management efforts on verifying that their model works as intended.

Model validation is a more challenging and important process that requires an organization to face up to some critical tradeoffs. At the most basic level, validation involves asking whether the model meets its users' requirements, and checking the quality of the data sources used to estimate its parameters. At a higher level, validation may involve a review of the process used to develop the model (e.g., the way domain experts were utilized, and the choice of quantitative or planning methodologies that were used). At a still higher level, validation can sometimes involve challenging questions about why one theoretical framework instead of another was used as the basis for a model (e.g., equilibrium versus non-equilibrium theories). At the highest level, validation involves confirming whether users and modelers are cognizant of the trade-offs they have made between a model that has a high fidelity to (i.e., fit with) historical data, a model with a high robustness to uncertainty, and a model that generates high confidence in its predictions. Unfortunately, this is not a trade-off that is well understood.

Models that very accurately predict historical data (i.e., that have very impressive backtesting results) are generally not very robust to uncertainty. Either parameter estimation errors or changes in the nature of the underlying system being modeled can generate very large variances between predicted and actual

outcome. An excellent example of this phenomenon is a simple asset allocation model that is based on historical returns, standard deviations and correlations between different asset classes. Such models always show impressive backtested portfolio returns; however, because future asset class returns, standard deviations and correlations tend to deviate from their historical values, the actual portfolio returns produced by a model are often poor, as many an investor can attest. More robust models generate smaller prediction errors for a given level of variability in their input parameters. Confidence in a model's predictions is increased when there is little difference in its predictions and the predictions produced by other equally credible models (which can be quantitative models, or the intuitive forecasts of experts). Unfortunately, as Ben-Haim and Hemez have shown, there is an unavoidable trade-off between these three aspects of model validation and quality, with increases in any two quality dimensions requiring a decrease in the third (see Yakov Ben-Haim and Francois Hemez, *Robustness, Fidelity and Prediction-Looseness of Models*, Proceedings of the Royal Society, A, 8 January 2012).

For example, a model may have high fidelity to historical data and may generate predictions that are very close to those produced by experts using intuitive means. However, this implies a relatively low level of robustness to uncertainty or lack of knowledge about input parameters and the form of the model itself. Alternatively, a model may have a high robustness to uncertainty, and high confidence in prediction, but have a low fidelity to historical data. Such models are often found in situations where rapid change is underway (e.g., financial market models since the 2008 crisis). Ultimately, the credibility of a model with its users depends on their awareness of these tradeoffs, and their agreement with the way they have been made by the modeler (which, obviously, brings us back to the quality of the model development process).

Conclusions

Models are critically important to business success, and thus an important but often overlooked source of enterprise risk. As such, model verification and validation are critical risk management activities. However, risk management also requires wisdom, and the ability to recognize the limitations of our efforts. Philosophically, it is logically impossible to completely verify and validate a model, as this would require a degree of omniscience that neither the modeler nor an independent assessor can possess. At best, one can say that a model has passed a set of verification and validation tests, and not been falsified by them, which should raise our confidence in its use, while accepting that the model itself is a simplified representation of a more complex and uncertain reality. As risk managers, we therefore also must focus on the resilience of our organizations, not just to the full range of outcomes predicted by our models, but also to the inevitable negative surprises that will sometimes result from those models' unavoidable shortcomings.

Tom Coyne consults on enterprise and investment risk management. He began his career as a credit officer at Chase Manhattan Bank. He subsequently spent almost twenty years as a management consultant, specializing in turnarounds and growth. He has also served as the CFO, CRO, and CEO of public and private companies. Tom is a frequent contributor to investment research publications, helping investors to make sense of uncertainty and reduce their exposure to severe losses. He provided subscribers with advanced warning of both the 2001 technology bubble collapse and the 2008 global financial crisis. He is also a member of the top ranked team in the Intelligence Advanced Research Projects Agency's multiyear forecasting tournament. Tom can be reached at tcoyne@sachuestadvisors.com