Super Models

By Timothy Paris

This article emerged from a series of recent presentations I gave about the development and maintenance of policyholder behavior models, the differences between models and assumptions, and how all of this can be used to quantifiably improve risk management. An important thread running through all of this is the ability to visualize and communicate highly technical concepts to colleagues and non-actuarial stakeholders. So while a certain amount of prose is inevitable, I have suppressed exhaustive numerical details and formulas in favor of a series of figures to illustrate how super models can help you and your company manage the risks in your business more effectively.

What is a super model? Of course, beauty is in the eye of the beholder, but I posit that in this context, super models are developed based on rigorous data analytics techniques, and they provide you with a range of potential outcomes, their financial impact and metrics that you can use to evaluate when material changes are necessary. "Assumptions" can be extracted from your super model for various applications, but the super model itself is more robust than that. It is a framework for analysis and risk management, not a point-in-time set of numbers.

WHAT IS THE FORM OF THE SUPER MODEL?

While my firm's particular focus is on annuity policyholder behavior models, the key underlying issues transcend product lines. In general, we are attempting to model the probability p of an event occurring, based on a function of a combination of factors $\vec{X}$ and coefficients $\vec{\beta}$:

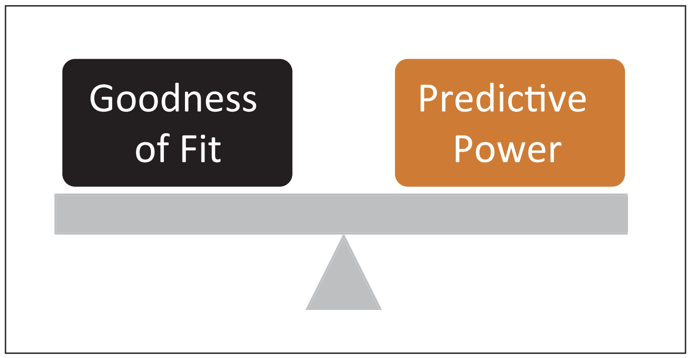
$$p = f(\vec{\beta}, \vec{X})$$

Admittedly, this is not much of a picture, yet the simplest equations are often the most beautiful. For example, we may wish to model the probability that a fixed indexed annuity contract makes a partial withdrawal in a given month, based on a combination of factors such as duration, the presence of a guaranteed lifetime income benefit, contract size, age and tax status, along with some interaction terms, as reflected in a generalized linear model. I find it remarkable when I observe companies that do not establish a baseline of functional form, and simply assemble a brittle set of numerical assumptions based on recently observed experience.

HOW WILL YOU SOLVE FOR FACTORS AND COEFFICIENTS?

Once the functional form is established, solving for the model factors and coefficients is a very challenging exercise, and there is typically a range of reasonable answers. Fundamentally, we are trying to build a model for something that will happen in the future. We typically calibrate such a model to some historical experience data, and test its predictive power against other data that is held out from that calibration process (see Figure 1). This often requires actuarial judgment and thoughtful trade-off decisions.

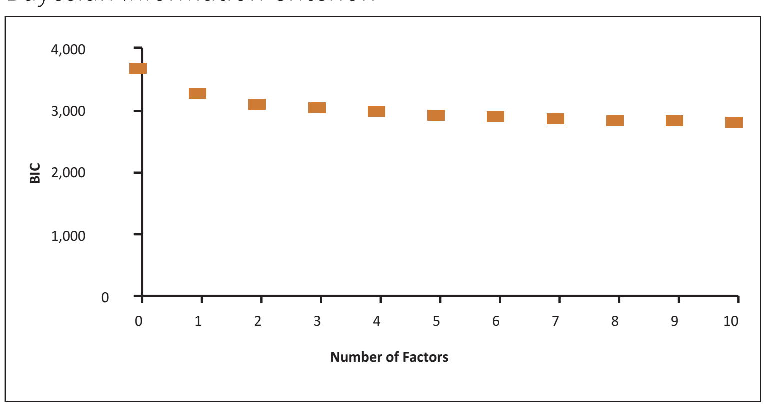
Figure 1
A Delicate Balance



MEASURING GOODNESS-OF-FIT

There are many ways to measure how well your model fits historical experience data, including metrics such as the Bayesian information criterion (BIC), as seen in Figure 2. A super model will often fit the historical experience very well using a relatively small number of factors that make business sense, sidestepping the pitfall of overfitting to noise.

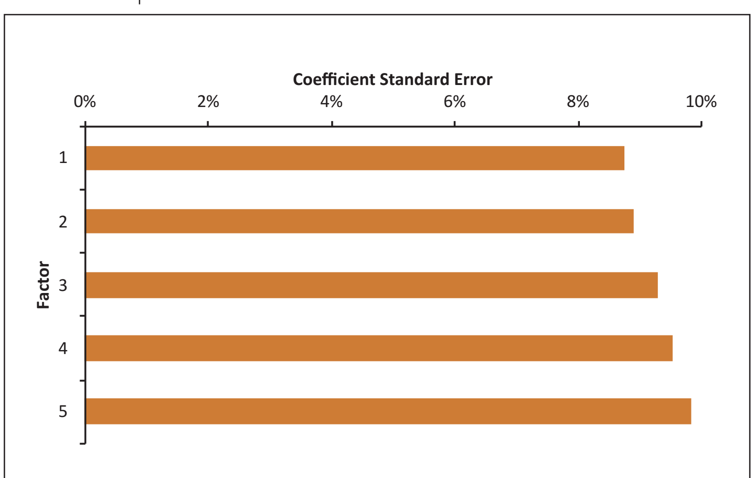
Figure 2
Bayesian Information Criterion



SUPER MODELS HAVE RANGE

Unlike mere assumptions, which are usually a defined set of numbers, sometimes quite elaborate-looking, that are often subject to endless seemingly arbitrary annual "unlocking," super models not only have baseline coefficient estimates for the model factors, but also standard error terms for each, in order to provide a sense of the range of possible outcomes based on historical data. By definition, no model is perfect, so super models attempt to quantify their own degrees of imperfection. This way, you are much better able to distinguish noise around modeled behavior from substantive changes. The pattern in Figure 3 is representative, with the most important factors having the lowest standard error terms, and hence higher confidence in the coefficient estimates.

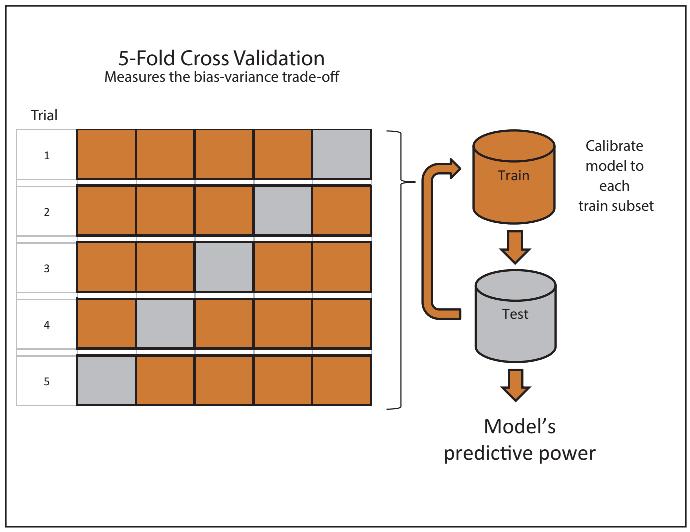
Figure 3
Example of Standard Error for Model Factors



SUPER MODELS HAVE WELL-TESTED PREDICTIVE POWER

While it is helpful to understand the model's goodness of fit to historical data, it is vitally important to quantify the model's predictive power *relative to data held out from the model development process*. This is essentially a sampling exercise, and many approaches can be insightful: simple splits like 60 percent of data for model development "training" and 40 percent for model "testing"; using the first several years of data to "predict" the last year; or cross-validation techniques like the one illustrated in Figure 4. On these bases, actual-to-expected ratios help you to determine which models perform better than others and what range of experience you may reasonably expect for the future.

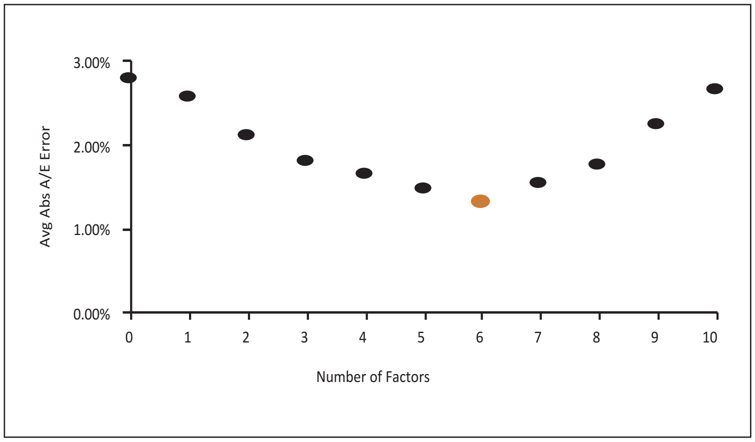
Figure 4
Cross-Validation Technique



KEEP IT SIMPLE

The more company data available to build your model, the greater the temptation to over-complicate. Initially, you will typically find that each additional explanatory factor you add to your model should improve its goodness of fit to historical data and its predictive power. So more is better—to a point. Goodness of fit tends to provide only diminishing returns with additional factors, and the improved fit to historical data is often just noise that may not be predictive of data held out from the model development, or of the future (see Figure 5). At some point you will need to employ actuarial judgment to determine when enough is enough. Ideally, this judgment will be guided by your company's objective risk management directives and actuarial governance.

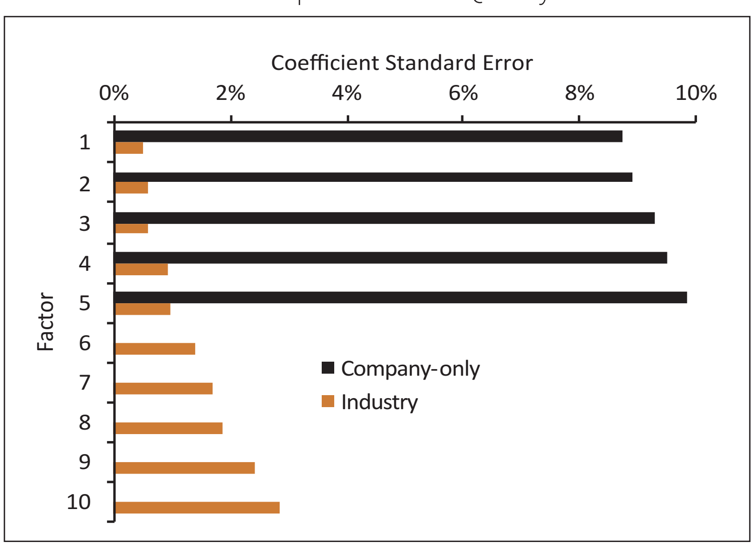
Figure 5
Diminishing Returns and Risks of Overfitting With Additional Factors



MORE DATA USUALLY BEATS MORE COMPLEX MODELS

On the other hand, if you are able to access additional relevant data to include in your model development process, such as data from external databases, industry experience studies, company affiliates, new business or reinsurers, complex models or models with many factors can often be statistically justified. Oftentimes, when limited to your own company's data, only a few model factors will be statistically justified—it is difficult to distinguish noise from real systemic effects. If you can access and use such external data, the quality of your model will tend to improve dramatically, as illustrated in the reduction in coefficient standard error terms in Figure 6 using industry data that is about 40 times larger than company-only data.

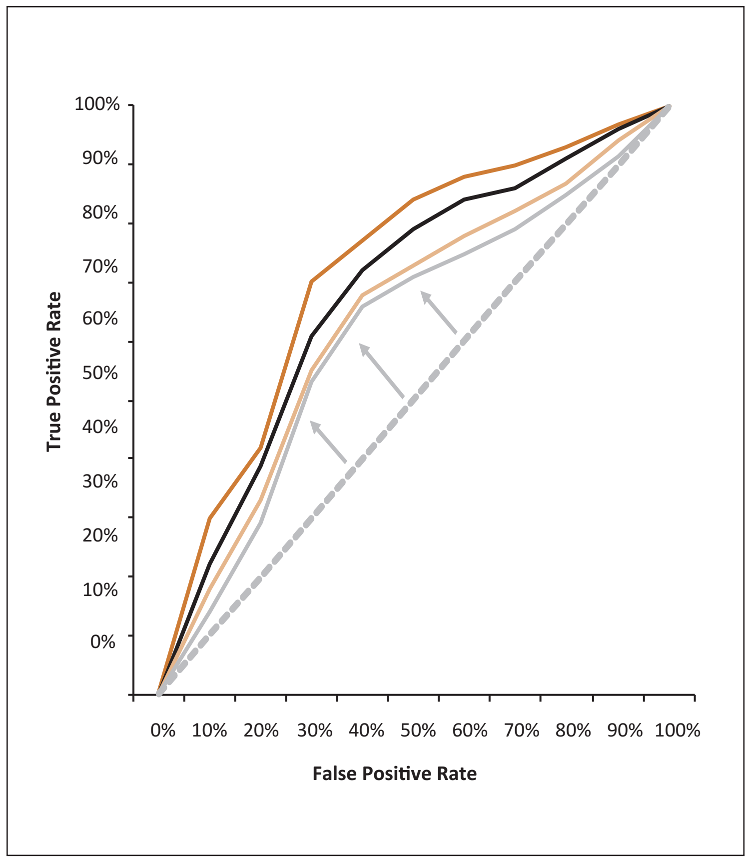
Figure 6
Additional Data Improves Model Quality



SUPER MODELS ARE SENSITIVE—AND SPECIFIC

Actuaries often use models to predict binary outcomes, such as whether withdrawals or deaths occur. Satisfactory aggregate model metrics are necessary but are not necessarily sufficient to qualify for super model status. We want a model that correctly predicts both of the possible binary outcomes. The statistical terms for these are sensitivity and specificity, and they are illustrated with the receiver operating characteristic (ROC) curve. Super models will have steep ROC curves, like those illustrated in Figure 7.

Figure 7
Receiver Operating Characteristic Curve



LOOK CLOSER

Continuing this theme, you should look closely at how well your model predicts important cohorts within the aggregate data, such as each of the modeled factors and any noteworthy factors that may not be explicitly included in the model. As illustrated in Figure 8 for one factor, super models tend to perform well at this level of granularity too, especially for the cohorts that comprise the bulk of the data as represented by the higher red dots in the center of the graph. This should give you confidence that even if your business mix changes along these dimensions, your super model will continue to look great.

Figure 8
Focus on Key Cohorts

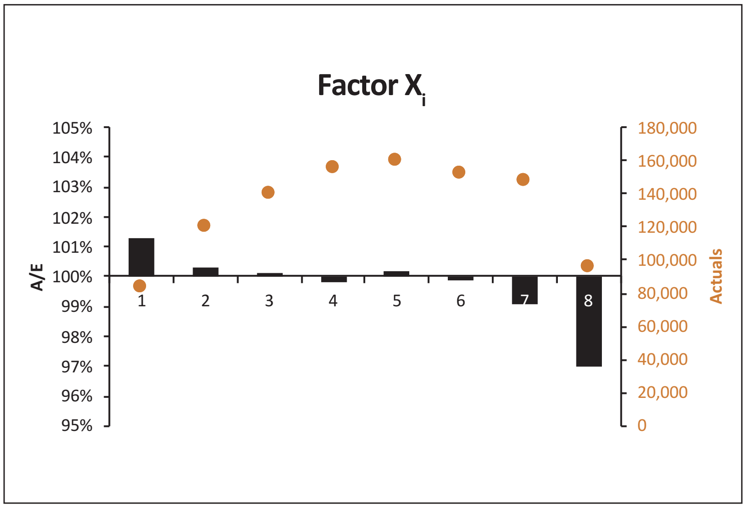


Figure 9
Lift Curve

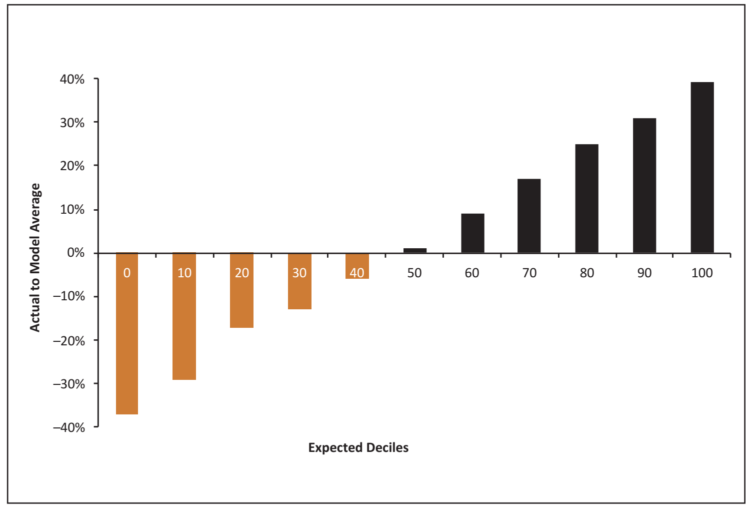
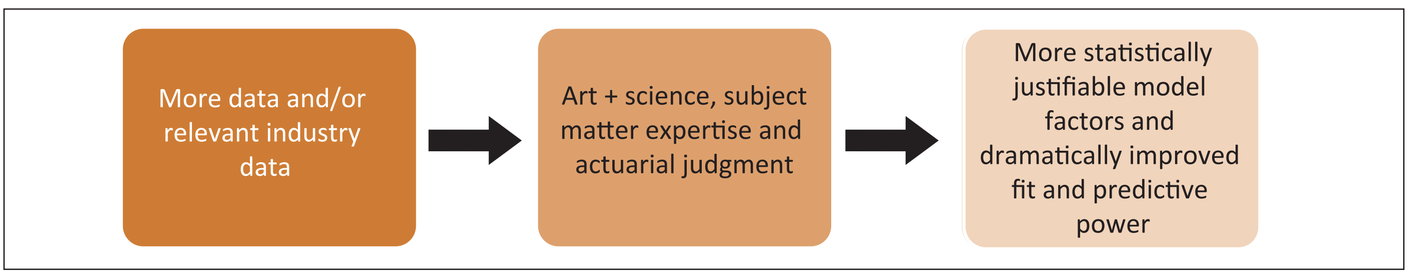


Figure 10
The Power of Super Models



SUPER MODELS GIVE YOU A LIFT

As part of your model validation process, you should find your super model also has the sensible property that, as the model's expected deciles increase, actual-to-average-expected values should also increase from negative to positive. This "lift curve" illustrated in Figure 9 is often accompanied by related metrics such as the Gini coefficient.

CONCLUSION

Actuarial super models exist. And they tend to be way better with more data, as described in my recent article in *The Actuary*.[1] I would venture that more and more actuarial super models are on the way, considering our increasing focus on this type of work. Regardless of the algorithms or software you use, their telltale characteristics are that they have a rationale for existence based on rigorous data analytics techniques, and they provide you with a range of potential outcomes, their financial impact and metrics you can use to evaluate when material changes are necessary. However, while there is a lot to like, you should not fall in love, since coefficients, factors and even the functional form of the super model itself will likely change. "Assumptions" can be extracted from your super model for various applications, but the super model itself is more robust than that. It is a framework for analysis and risk management, not a point-in-time set of numbers. So don't settle for less. And if you want your stakeholders to understand this, pictures of super models can be really helpful (see Figure 10). ■

Timothy Paris, FSA, MAAA, is chief executive officer at Ruark Consulting LLC and is also the leader of the new Assumption Development and Governance Subgroup of the Modeling Section. He can be reached at *timothyparis@ruark.co.*

ENDNOTE

1 Paris, Timothy. 2018. When is Your Own Data Not Enough? How Using External Data can Strengthen Results. *The Actuary* 15, no. 3:28–33, *http://theactuarymagazine.org/wp-content/uploads/2018/06/act-2018-vol15-iss3.pdf.*