

Simple versus complex forecasting: The evidence

Kesten C. Green, and J. Scott Armstrong

Working Paper 224, January 2015

This paper was published as

Green, K. C. & Armstrong, J.S. (2015). Simple versus complex forecasting: The evidence, [Journal of Business Research](#), 68, 1678–1685.

Acknowledgements: Kay A. Armstrong, Ron Berman, John E. Boylan, Peter Fader, Antonio García-Ferrer, Robert Fildes, Paul Goodwin, Wilpen Gorr, Andreas Graefe, Robin M. Hogarth, Xueming Luo, Philip Stern, Aris Syntetos, Christophe van den Bulte, and Malcolm Wright provided helpful reviews. We are grateful to the many authors we cite who replied to our appeal for help, suggested improvements to our summaries of their research, and pointed us to additional evidence. Hester Green, Emma Hong, Sue Jia, Yifei Pan, and Lynn Selhat edited the paper. The authors of this article are responsible for any remaining errors or omissions.

Contact information: Kesten C. Green, University of South Australia Business School, and Ehrenberg-Bass Institute, GPO Box 2471, Adelaide, SA 5064, Australia; kesten.green@unisa.edu.au. J. Scott Armstrong, The Wharton School, University of Pennsylvania, 700 Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, U.S.A., and Ehrenberg-Bass Institute, Adelaide; armstrong@wharton.upenn.edu.

Abstract

This article introduces the Special Issue on simple versus complex methods in forecasting. Simplicity in forecasting requires that (1) method, (2) representation of cumulative knowledge, (3) relationships in models, and (4) relationships among models, forecasts, and decisions are *all* sufficiently uncomplicated as to be easily understood by decision-makers. Our review of studies comparing simple and complex methods—including those in this special issue—found 97 comparisons in 32 papers. None of the papers provide a balance of evidence that complexity improves forecast accuracy. Complexity *increases* forecast error by 27 percent on average in the 25 papers with quantitative comparisons. The finding is consistent with prior research to identify valid forecasting methods: all 22 previously identified evidence-based forecasting procedures are simple. Nevertheless, complexity remains popular among researchers, forecasters, and clients. Some evidence suggests that the popularity of complexity may be due to incentives: (1) researchers are rewarded for publishing in highly ranked journals, which favor complexity; (2) forecasters can use complex methods to provide forecasts that support decision-makers' plans; and (3) forecasters' clients may be reassured by incomprehensibility. Clients who prefer accuracy should accept forecasts *only* from simple evidence-based procedures. They can rate the simplicity of forecasters' procedures using the questionnaire at simple-forecasting.com.

Keywords: analytics; big data; decision-making; decomposition; econometrics; Occam's razor.

Introduction

This article provides an introduction to this Special Issue on simplicity in forecasting. The call for papers was subtitled “Conditions and complexity in forecasting,” and the objective was to publish “research to improve forecasting knowledge by comparing the usefulness of simple and complex alternatives under different conditions.”

A trend toward complex forecasting has been underway for the past half-century or more. Econometricians who believe that complex statistical procedures yield greater forecast accuracy have led the trend (see, e.g., Armstrong 1978). The trend is at odds with the common belief among scientists that scientists should strive for simplicity. The preference for simplicity in science can be traced back to Aristotle (Charlesworth, 1956), and is commonly identified with the 14th Century formulation, *Occam’s razor*. Indeed “since that time, it has been accepted as a methodological rule in many sciences to try to develop simple models” (Jensen 2001, p. 282). Zellner (2001) concludes social scientists too should strive for simplicity. He was joined in this conclusion by the 21 authors contributing to the book, *Simplicity, Inference and Modelling* (Zellner, Keizenkamp, and McAleer, 2001).

This article first draws upon prior literature to develop an operational definition of simplicity in forecasting, then uses the definition to identify and analyze comparative studies that could be expected to provide evidence on the conditions under which complexity is useful. The review of studies includes new evidence presented in this Special Issue. Finally, this article examines evidence on why researchers, forecasters, and decision-makers are, despite the theoretical and empirical objections, attracted to complexity.

Defining simplicity in operational terms

Simplicity in forecasting seems easy to recognize, yet is difficult to define. The first definition in the *Oxford English Dictionary’s* OED Online (2014) is, nevertheless, a useful starting point: “The state or quality of being simple in form, structure, etc.; absence of compositeness, complexity, or intricacy.”

For the purpose of making practical distinctions between simple and complex forecasting, this article defines simple forecasting as processes that are understandable to forecast users. Specifically, the forecasting process must be understandable with respect to methods, representation of prior knowledge in models, relationships among the model elements, and relationships among models, forecasts, and decisions.

Complexity in forecasting is the opposite of simplicity. In contrast to some discussions of complexity in forecasting, by our definition complexity is *not* a function of the number of variables. Nor is complexity a function of the effort required to develop a model.

To conclude whether or not an instance of forecasting is simple, as defined here, ask forecast users if they understand—and, if so to explain—the forecasting method, how the specific model represents prior knowledge, how any parts the model has are related to each other, and how and why a forecast from the model would help them to make a better decision. A structured questionnaire to derive a measure of the simplicity of the forecasting procedures from forecast users’ understanding—the Forecasting Simplicity Questionnaire—is available from simple-forecasting.com.

The test of simplicity provided by the questionnaire has face validity. Recounting his correspondence with Nobel Laureates and other leading economists, Zellner reports James Tobin telling him that he and his Council of Economic Advisors colleagues were skeptical of complex models of the economy because they “could not understand the workings and outputs of such models, and thus did not have much confidence in them” (Zellner 2001, pp. 243-244).

Zellner (2001, p. 242) observes, “Some years ago, I came upon the phrase used in industry, ‘Keep it simple stupid’, that is, KISS, and thought about it in relation to scientific model-building.

Since some simple models are stupid, I decided to reinterpret KISS to mean ‘Keep it sophisticatedly simple.’” With that distinction in mind, this article is concerned primarily with comparisons of complex forecasting with simple forecasting procedures that have been validated by experimental comparisons.

Simple methods

Simple forecasting methods are relatively few compared to complex methods, which are limited in number only by the imaginations of statisticians. The titles and abstracts of forecasting papers in academic journals attest to the proliferation of complex methods. Not only managers, but also practitioners and many researchers are likely to struggle to comprehend typical forecasting papers.

Incomprehension of forecasting methods, even by the people who pay for them, seems common. For example, as part of a three-person consulting team, the second author of this article interviewed several analysts in a large firm to assess their understanding of a complex model provided at high cost by an outside vendor. The model was designed to forecast the effects of advertising expenditures on the company’s market share. The vendor provided courses to explain the method to their clients. Even so, none of the analysts could explain how the model worked (Armstrong and Shapiro, 1974).

Simple representation of prior knowledge in models

A scientific, or evidence-based, approach to forecasting requires an effort to represent cumulative knowledge (Armstrong, Green, and Graefe, this issue). Before the 1970s, econometricians often based their forecasting models on a priori analyses. They used domain experts’ knowledge and what evidence they could glean from prior research to guide their selection of variables, to determine directions and the nature of the relationships, and to estimate the magnitudes of the relationships. While the process is a logical scientific procedure and is simple to explain, much time and effort by experts is often required in order to carry it out.

In contrast to the high cost of a thorough a priori analysis, applying complex statistical methods to large databases is inexpensive. McCloskey and Ziliak (1996) and Ziliak and McCloskey (2004) show that many researchers follow the low-cost approach. Their analyses of *American Economic Review* papers found that 75 percent of the papers in the 1980s that used regression analysis went beyond statistical significance to consider other information when selecting variables for regression models. The figure dropped to 32 percent in the 1990s.

Regression analysis identifies statistical patterns in a particular set of data. If the data are non-experimental, no matter how “big” they are, there is little reason to expect the process to identify causal relationships (Armstrong 2012; Armstrong, Green, and Graefe, this issue). In practice, a big data set is likely to include variables that are not independent of one another, variables that vary little or not at all, and irrelevant variables, while excluding variables that are important. The need for theory, domain knowledge, experimental data, and careful thinking for specifying and estimating causal models has not changed.

Bayes’ method provides another way to incorporate prior knowledge in forecasting models. The method has the disadvantage of being too complex for most people to understand. We have been unable to find evidence that Bayesian approaches yield ex ante forecasts that are more accurate than forecasts from simple evidence-based methods. The first M-Competition (Makridakis, Anderson, Carbone, Fildes, Hibon, Lewandowski, Newton, Parzen, and Winkler, 1982) includes tests of Bayesian forecasting for 1 to 18 period ahead forecasts for 997 time series. Forecasts from simple methods, including naïve forecasts on deseasonalized data, were more accurate than Bayesian forecasts on the basis of mean absolute percentage error (MAPE). Forecasts from the benchmark deseasonalized single exponential smoothing method reduced error by 12.4 percent (from Makridakis et al., 1982, Table 2a). Bayesian forecasts were not included in subsequent M competitions. Graefe, Küchenhoff, Stierle, and Riedl (2014) found that simply averaging forecasts from different methods yields forecasts that reduced error by an average of 5 percent across five

studies compared to those from Bayesian approaches to combining economic and political forecasts. Goodwin (this issue) demonstrates that for many forecasting problems that involve choosing between two alternatives, two simple methods would each lead to the same decision as Bayes' method.

The simplest representations of prior knowledge in forecasting models are no-change models. Forecasts from appropriately formulated no-change models are hard to beat in many forecasting situations, either because prior knowledge is insufficient to improve on no-change or because prior knowledge leads to the conclusion that the situation is stable.

Simple relationships among the model elements

Decomposition provides a path to simplicity for many forecasting problems. Decomposition in forecasting consists of breaking down or separating a complex problem into simpler elements before forecasting each element. The forecasts of the elements are then combined.

Some researchers suggest that decomposition increases complexity relative to forecasting the aggregate directly—such as the works cited by Brighton and Gigerenzer in this issue—but that is not the case with the definition proposed in this article.

Decomposition is a key strategy for simplifying problems in management science, and in other scientific fields. Decomposition can be used with any forecasting method. The method is most useful when different elements of the forecasting problem are forecast by different methods, when there is valid and reliable information about each element, the elements are subject to different causal forces, and when they are easier to predict than the whole.

A study on forecasting traffic accidents by García-Ferrer, de Juan, and Poncela (2006) provides evidence on the benefits of disaggregation when these conditions are met. Their approach of disaggregation by estimating separate models for urban and other roads produced forecasts that were more accurate for between 76 and 85 percent of the 63 comparisons, depending on the criterion used.

If there are few data on each element, however, decomposition may not improve forecast accuracy. Huddleston, Porter, and Brown (this issue) examine the trade-off in their tests of different approaches to forecasting highly variable district-level burglary rates.

The relationships among the elements of the decomposed problem should be simple. Decomposition based on additive relationships, an approach that is often referred to as segmentation, is ideal. Decomposition based on multiplicative relationships—in which the elements are multiplied together to obtain a forecast of the whole—is somewhat more complex, carrying the risk that errors will multiply; nevertheless, multiplicative decomposition is often useful for simplifying complex problems.

Many ways are available to decompose forecasting problems. A common approach to forecasting sales, for example, is to forecast market size and market share separately. Another approach is to decompose a time-series by estimating the starting level—“nowcasting”—and forecasting the trend separately. Combining nowcasting with trend forecasting is an old idea that does not appear to be widely used, and comparative tests are few. Nevertheless, the two studies described in Tessier and Armstrong (this issue) suggest that substantial error reduction is possible.

Transforming variables can help to avoid complexity in a model. Perhaps the most common approach is to transform multiplicative relationships into additive relationships by using logarithms. The coefficients of variables in logarithmic form are known as elasticities. Elasticities represent relationships in an easily understood and useful way; they are the expected percentage change in the variable being forecast arising from a one percent change in the causal variable. The intuitiveness of elasticities allows clients to readily transform their knowledge into expectations about the magnitude of causal relationships.

The index method is an approach to decomposition that is appropriate for situations with many causal variables. The index method involves identifying and examining each causal relationship individually before combining them in a forecasting model. This easily understood approach avoids

the complications that arise from using regression analysis to develop a forecasting model. As Armstrong, Green, and Graefe (this issue) describe, index models provide forecasts of election results that are substantially more accurate than those from multiple regression models.

Simple relationships among models, forecasts, and decisions

The relationships among the models, forecasts, and decisions need to be clear to decision-makers in order to help them to choose among alternative courses of action. One way to achieve that clarity is to describe the proposed method, and then ask the decision-makers what decisions they would make in response to different hypothetical forecasts. The forecasts should include forecasts of costs and benefits, and likelihoods. Hogarth and Soyer (this issue) found that forecast users make better use of uncertainty information about forecasts when they are able to observe possible outcomes with the aid of simulation software, than when they are provided with standard statistical information about the model.

Indeed, the complex statistics typically provided with regression models are unlikely to help decision-makers to make better decisions, as they confuse even statisticians. Soyer and Hogarth (2012) asked 90 economists from leading universities to interpret standard regression analysis summaries. Roughly two-thirds of their answers to three relevant questions were substantively wrong.

Regression statistics can divert attention from the decision-maker's need to assess the effects of causal variables. Regrettably, the attention of decision-makers is commonly subject to diversion of that kind. Academics are not immune, as Ziliak and McCloskey (2008) show with many examples from econometric forecasting.

The R^2 statistic continues to mislead many analysts and decision-makers, despite repeated warnings that the statistic is a poor measure of predictive ability. Armstrong (1970) demonstrates that even with data that are random numbers, a high R^2 is easily achieved by using stepwise regression in combination with other common exploratory data analysis techniques. Armstrong (2001, p. 461) identifies six studies on the use of R^2 , and finds little relationship with forecast accuracy. Similarly, Peach and Webb (1983) estimate 50 econometric models in each of three standard mathematical forms using 95 and 134 observations. The models involved three independent variables and one dependent variable chosen at random from the National Bureau of Economic Research's data bank. The resulting R^2 and t statistics were similar to those of established models published in the economics literature. Efforts by forecasters to improve R^2 are likely to harm predictive validity because each new model specification leads the modeler away from the original theoretical formulation, assuming there was one.

Ziliak and McCloskey (2008) illustrate the harm caused by statistical significance testing with examples taken from across the sciences. Cumming (2012) describes additional examples. Much of the harm caused by tests of statistical significance arises because they divert attention from the pursuit of important information, such as the likely costs and benefits of a proposed policy. For example, Hauer (2004) reports that the use of statistical significance led to poor decisions on automobile traffic safety policy, such as the right-turn-on-red rule. Forecasts of the effects of the rule on accidents and deaths, and the time saved by drivers, would have been more useful than the results of statistical significance tests.

Simpler and more useful measures are available to forecasters than R^2 and statistical significance. For example, to assess which forecasting method is best, the relative absolute error (RAE)—being the size of the forecast error relative to the size of the forecast error from a relevant no-change model—is useful and simple to understand. For production and inventory control decisions, the mean absolute error (MAE) is a simple and useful measure (Armstrong and Collopy, 1992), although it can mislead if demand is intermittent (Boylan and Syntetos 2006; Teunter and Duncan 2009).

Effects of complexity of methods and models on accuracy

Simplicity in forecasting has the obvious advantage of encouraging engagement and criticism by facilitating understanding. In addition, simplicity aids in detecting mistakes, important omissions, ludicrous variables, unsupported conclusions, and fraud. But how accurate are simple methods?

To answer that question, the authors of this article searched for studies that compare the accuracy of forecasts from simple versus complex forecasting procedures. To do so, they used keywords to search for papers on the Internet, examined references in key studies, and contacted key authors. The emphasis of the search was on finding studies showing that, under certain conditions, complex methods provide forecasts that are more accurate. To help ensure that accurate representation of the studies identified in the search and reduce the risk of overlooking important studies, the authors attempted to contact the living authors of papers cited in substantive ways in this article. Twenty-one out of the 27 authors with known email addresses responded. Their replies led to improvements in the paper. The authors of one paper disagreed with the analysis of their findings on forecasts from neural networks presented in this article. Details of the analysis are available at simple-forecasting.com, along with a statement by the two authors of their own conclusions about their study.

When this article was near completion, the authors sent drafts to email lists with requests for further evidence, especially evidence that conflicts with the article's conclusions. Responses provided opinions for and against the conclusions along with a few references to relevant studies. The request led to few additions to the evidence because the studies proposed by respondents did not provide evidence on complexity versus simplicity as the concepts are defined in this article. Typically they associated complexity with the number of variables, a factor explicitly excluded from this article's definition.

The following sections review the evidence on simplicity versus complexity for judgmental, extrapolative, and causal methods—and for combining forecasts. In the judgement of the authors of this article, the differences in the complexity of the methods in the comparative studies identified were typically large, though the simpler methods were often not as sophisticatedly simple as they might have been. In other words, the comparisons we identify may well *understate* the relative difference in performance between complex and simple evidence-based methods. Full disclosure of ratings is provided at simple-forecasting.com so that readers can make their own judgments.

Judgmental methods

Humans' capacity for mental processing of information has modest limits. As a consequence, judgmental forecasters rapidly reach a point beyond which further information does not help them to make more accurate forecasts. Moreover, humans, no matter how clever, are unable to learn about complex relationships from experience alone. Thus, without structured methods, we are all ill equipped to make forecasts about complex, uncertain situations.

The structure, however, need not be complex. In a study on forecasting demand for a new form of transportation, a quasi-experiment compares complex and simple methods for obtaining data on purchase intentions. Subjects in the complex treatment visited a product clinic that allowed them to sit in a prototype car, see descriptive wall posters, watch a descriptive movie, and participate in a focus group. Then they completed an intentions survey. Another group was mailed a two-page description of the system, the second page being a picture of the prototype car, and participants completed the same intentions survey. The reported intentions were similar for the two groups (Armstrong and Overton, 1971).

Lees and Wright (2004) provide further evidence on the effects of additional information on intentions surveys. They obtained purchase intentions for five diverse product concepts. The proposed products were presented in one of three forms: simple factual description, extended promotional description, and extended promotional description with artwork. They obtained responses from between 565 and 625 respondents per treatment, and find little difference in

intentions to purchase between those who had been given the simple descriptions and those given more complex descriptions.

Simulated interaction provides a simple way to structure comprehensive information about complex situations that involve interactions between parties whose interests diverge. The method involves asking people to take on the roles of key participants, providing them with a short description of the situation, and leaving them to interact in ways that are appropriate to the situation. The typical decision reached in a simulated interaction is used as the forecast. Green (2005) obtained 105 simulated interaction forecasts of decisions in eight conflict situations. The forecasts reduced error by 47 percent compared to the 101 forecasts by experts on the complex method of game theory

The interactions in the simulated interaction method appear to enable people to make better use of more information about complex situations than is the case with unaided judgment. Green and Armstrong (2011) test that assumption by obtaining 101 role-thinking forecasts of the decisions that would be made in nine diverse conflicts from 27 Naval postgraduate students and 107 role-thinking forecasts from 103 novices. Given the complex task of thinking about the roles, objectives, strategies, and interactions of the parties in conflict situations, the accuracy of both groups' forecasts was little better than the 28 percent correct that could be expected from guessing. Thus, neither experts nor novices were able to make good use of information about complex situations by thinking hard about them, whereas the same information could be simply and realistically modeled by simulated interaction, the use of which reduced forecast error by 41 percent for the nine situations.

Extrapolation methods

Extrapolation methods that incorporate more data are likely to improve forecast accuracy. That said, more recent data are typically more relevant, especially for short-term forecasts. To address that dilemma, Brown (1956) proposes exponential smoothing. Exponential smoothing forecasts turned out to be more accurate than those from commonly used methods such as judgmental extrapolation and moving averages, and the additional complexity that arises from using exponential smoothing is trivial.

A second sophisticatedly simple improvement to extrapolation involves damping the trend in an exponential smoothing model toward zero when there is uncertainty (see Gardner 2006 for a review). By our definition, damping only trivially increases complexity. All is done automatically via decomposition, along with additions and multiplications.

A third sophisticatedly simple improvement to extrapolation is achieved by making adjustments for seasonality when the interval of forecasts is shorter than one year. The now widespread use of seasonal factors is largely influenced by the work of Shiskin (1965). Seasonal adjustment is a form of decomposition, and can be implemented using addition or multiplication. Seasonal adjustments can reduce forecast errors substantially. For example, for 68 monthly economic series from the M-Competition, Makridakis et al. (1982, Table 14) found that seasonal adjustments reduced the MAPE of forecasts for horizons out to 18 months from 23.0 to 17.7 percent—a 23 percent error reduction. In the M2 Competition, seasonal adjustment reduced MAPEs for 23 monthly series forecast for up to 15 months ahead by 41 percent (from Exhibit 1, Makridakis, Chatfield, Hibon, Lawrence, Mills, Ord, and Simmons, 1993).

The simple no-change extrapolation model is a strong competitor in many forecasting situations, often after the data have been decomposed or adjusted—e.g., by or for price level or population. The model is usually formulated as “no change from the current level” but sometimes as “no change from the long-term trend.” A well-known demonstration of the power of the no-change model is the “random walk down Wall Street” (Malkiel, 2012). Random walk is an economists' term to describe the behavior of a time series without a predictable pattern, and hence the next value in the series is expected to be the same as the previous one. Researchers have been unable to improve upon the no-change model for forecasting day-ahead prices in the stock market.

Similarly, Schnaars and Bavuso (1986) compare the accuracy of 180 forecasts from the no-change model with 180 forecasts from each of six more complex extrapolation methods applied to each of 15 weekly economic indicators. The indicators included production, unemployment claim, and resource price series. On average, the no-change model yielded the most-accurate forecasts. The forecast errors, MAPEs, from the no-change model were half those of forecasts from the most complex extrapolation method tested, generalized adaptive filtering.

Over the past seven decades or so, authors of journal articles have proposed many complex extrapolation procedures. An early review of comparative studies suggests they have not led to improvements in the accuracy of forecasts, with 28 comparisons finding forecasts from the simpler method were as or more accurate than those from the more complex method and only 11 finding better accuracy from the more complex method (Armstrong, 1984).

Research since the 1984 review provides additional evidence. Smith's (1997) review found six studies in which extrapolations of population from complex models were no more accurate than those from simpler models, and only one study in which complex models were more accurate. The simple method was not, however, sophisticatedly simple.

Studies in marketing have shown similar results. Schnaars (1984) and Meade and Islam (2001) found that forecasts from complex curve-fitting models were no more accurate than those from simple extrapolation models. And in their review of evidence on methods for forecasting the trial of new consumer packaged goods, Fader and Hardie (2001) report that extrapolations using simple models—which were based on estimates of propensities to buy from early purchase data—provide more accurate forecasts than those from complex models estimated from the same data in two comparisons. Brighton and Gigerenzer (this issue) describe findings that experts' simple rules provide forecasts of customer behavior that are more accurate than mathematically sophisticated models advocated by researchers.

The M-Competition found that the simplest extrapolation methods suitable for the data used in the competition—deseasonalized no-change and single exponential smoothing—provided forecasts that were at least as accurate than those from all 16 of the more complex methods. On average, the two simplest methods provided forecasts for 1 to 18 months ahead for 1,001 time series that reduced MAPE by 34 percent compared to the forecasts from the more complex methods (from Table 2a, Makridakis et al., 1982). In the M2 Competition, the deseasonalized no-change and combined exponential smoothing forecasts reduced MAPE by 27 percent on average compared to forecasts from seven more-complex methods; namely five expert forecasters who had access to causal information, Box-Jenkins, and an autoregressive model (Exhibit 1 from Makridakis et al., 1993). In the case of the M3 Competition, the deseasonalized no-change, combined exponential smoothing, and Rule-Based Forecasting forecasts reduced MAPE by nearly 1 percent on average compared to forecasts from 17 more complex methods (Table 6 from Makridakis and Hibon, 2000). The relatively modest error reduction from the simpler methods in the M3 Competition presumably arises because the contestants learnt that naïve forecasts, exponential smoothing, and damping are hard to beat and, as a result, more contestants entered forecasts drawn from these methods via expert systems and proprietary software, and fewer of the complex forecasting methods that have failed badly were entered.

Forecasts from neural networks, a complex method, were 3.4 percent less accurate than relatively simple damped-trend forecasts in a test against 3,003 series in the M3-Competition (Makridakis and Hibon, 2000). In Crone, Hibon, and Nikolopoulos's (2011) subsequent forecasting competition, competitors entered forecasts from 27 methods—22 of which were neural network methods—for 1 to 18 month horizons for either 11 or 111 monthly time series. The competition organizers included forecasts from six methods that are simple using this article's definition. They are the naïve or no-change model without seasonal adjustment, seasonal adjustment, single exponential smoothing, Holt's exponential smoothing, dampened exponential smoothing, and a simple average of the exponential smoothing forecasts. The median RAEs of forecasts from the six simple methods are 9.9 percent smaller than the median RAEs of the forecasts from the 22 neural

network methods and 8.5 percent smaller than the median RAEs of the forecasts from all 27 of the complex methods entered in the competition when compared using geometric means. Moreover, the typical neural network forecast was four percent less accurate than forecasts from the naïve (no-change) model.

For another example, consider the task of forecasting intermittent demand. Syntetos, Babai and Gardner (this issue) compare forecasts from simple, no trend, exponential smoothing and two other simple techniques developed specifically for intermittent demand forecasts, with those from a complex method reported by Willemain, Smart, and Schwarz (2004). They find little difference in the forecasts for jewelry sales, and forecasts from the complex method were not as useful as those from simple exponential smoothing for electronics sales.

Nikolopoulos, Goodwin, Patelis, and Assimakopoulos (2007) compare methods for forecasting audience shares for a holdout sample of 12 TV sporting events. A simple average of the shares obtained by the five most analogous TV shows, from among 34 previous shows, reduced error by 31 percent compared to the average errors of forecasts from two multivariate regression and three neural network models. Their forecasts based on three analogies and one analogy are not as accurate as those based on five, which is consistent with the findings of Green and Armstrong (2007), and suggests why Brighton and Gigerenzer's (this issue) "single nearest neighbor" analogy forecasts perform relatively poorly in their test of methods for picking the city with the larger population of German city pairs: they used only a single analogy.

In another example of extrapolating from analogous data, Wright and Stern (this issue) compare the accuracy of sales forecasts for very different new products. The simpler, more intuitive method based on the sales of analogous products provided forecasts that were substantially more accurate—error was reduced by 43 percent—than the forecasts from three established complex models. Wright and Stern note that the established complex models are, in turn, simpler than four other models that were rejected in prior research due to the relative inaccuracy of their forecasts.

Causal methods

Causality is often complex. As a consequence, people often assume that complex methods and models will be needed for forecasting. Lesser (1968) observes that econometricians strive for complexity by using more equations, more complex functional forms, and more complex interactions in their regression models. Consistent with this, a survey of leading econometricians by Armstrong (1978) finds that the great majority of the 21 respondents agree with the proposition that more complex econometric methods will provide more accurate forecasts. A study on the accuracy of short-term economic forecasts by the National Bureau of Economic Research in the early 1970s concludes that, "the record reveals no clear-cut and sustained advantage of complex... forecasting systems" (Juster 1972, p. 23).

As early as the mid-1900s, econometricians proposed that simultaneous equations—which are complicated because they involve interactions among the equations—should lead to more accurate forecasts than simpler approaches. Armstrong (1985) found five comparative studies; the use of simultaneous equations failed to improve accuracy in any of them (p. 200).

Nikolopoulos, Goodwin, Patelis, and Assimakopoulos (2007) find that a simple one-causal-variable regression model, estimated from 34 observations, reduces the error of audience share forecasts by 37 percent, compared to forecasts from two more-complex regression models using the same data. One of the complex models incorporated all three available causal variables and the other was a stepwise regression model incorporating the best two variables, based on correlations. The simple model also provided an error reduction of 22 percent compared to the average of the forecast errors from three neural network models.

Fildes, Wei, and Ismail (2011) compared conditional one, two, and three-year forecasts of air passenger traffic flows from ten complex econometric models with exponential smoothing forecasts. Compared to the average complex econometric forecast, exponential smoothing forecasts reduced error across all horizons and overall by 10.8 percent on the basis of geometric RAE (from Fildes,

Wei, and Ismail, 2011, Table 6). Despite the fact that the econometric models use more information, the complexity penalty led to increased forecast errors.

The case for simple econometric methods became stronger when evidence on the value of equal-weights began to appear in the 1970s (e.g. Schmidt, 1971). Empirical evidence shows that simply assigning equal-weights to standardized predictor variables in a linear model usually yields ex ante forecasts that are at least as accurate as those from methods, such as regression analysis, that calculate optimal fits with the estimation data (see Graefe, this issue, for a review of the evidence).

Woike, Hoffrage, and Petty (this issue) provide further evidence on the advantages of uncomplicated weighting methods in their simulations of venture capital investment decisions. Their simpler weighting schemes provided predictions that led to more profitable decisions in most environments and provided the most reliable predictions when the environment was uncertain.

The Dean of North Dakota College of Pharmacy asked Gorr, Nagin, and Szczypula (1994) to develop a better model for forecasting prospective students' final GPAs. The college used a model consisting of a simple judgmentally weighted index of seven variables. Alternative models based on multiple regression, stepwise regression, and artificial neural networks provided no meaningful improvements in accuracy.

Graefe (this issue) examines evidence from U.S. presidential election forecasting. The ex ante forecast errors from versions of nine established regression models that equally weighted the causal variables were five percent smaller than those from the original models. In addition, Graefe demonstrates one of the major advantages of a simple additive equal-weights approach: the ability to include all variables that are important in a causal model. The error of ex ante forecasts from an equal-weights model that included all of the 27 unique causal variables from the nine original models was 48 percent lower than the error of the typical model and 29 percent lower than the error of the forecasts from the most accurate regression model.

Combining forecasts

Combining forecasts that incorporate different data and knowledge is a simple and easily understood way to represent prior knowledge. In a point of departure with Brighton and Gigerenzer (this issue), we regard combining as being an inherently simple forecasting procedure in keeping with the definition proposed in this article.

Complex methods for averaging have been proposed, but have not been met with success. An examination of Clemen's (1989) review of 209 studies on combining forecasts suggests that complex combining schemes cannot be relied upon to provide forecasts that are more accurate than those from simple averages. Studies since 1989 support Clemen's conclusion. For example, Duncan, Gorr, and Szczypula (2001, p. 209 Exhibit 6) find that using a complex method for combining forecasts of school revenues from a complex forecasting method *increases* forecast errors from 5.5 to 10.7 percent, or by 94 percent.

Lyon, Wintle, and Burgman (this issue) elicit confidence intervals for estimates of 311 quantities from 264 participants in 15 experiments. They test 13 complex approaches to combining the subjects' confidence intervals, and find that simple trimmed-means provide the most accurate point forecasts. Graefe, Küchenhoff, Stierle, and Riedl (2014) find that for economic forecasting and election forecasting, the simple average provided more accurate predictions than did a Bayesian approach to combining in five comparisons, and somewhat less accurate forecasts in one comparison.

Fildes and Petropoulos (this issue) provide evidence on forecasting method selection for combinations. Their findings support differential weighting in situations when there is prior evidence on which methods provide forecasts that are most accurate given the conditions.

Why simplicity repels and complexity lures

For what reasons do forecasters avoid simplicity? One is that if the method is intuitive, reasonable, and simple, would-be clients might prefer to do their own forecasting.

Another reason is that complexity is often persuasive. The “Dr. Fox study” found that university faculty and staff gave high ratings to a complex lecture, even though the content was nonsensical. Respondents commented that while they did not understand everything Dr. Fox said, he certainly “knew his stuff” (Naftulin, Ware, and Donnelly 1973). An extension by Armstrong (1980) describes tests using simple and complex versions of papers with identical content. Academicians rated the authors of research papers more highly when the papers were written in more complex ways. For additional experimental evidence, see Armstrong (2010, pp. 183–184).

Eriksson’s (2012) experiment provides additional evidence on the persuasiveness of complexity. The experiment included showing abstracts of two published papers to 200 subjects, all of whom were familiar with reading research reports and had post-graduate degrees. One of the abstracts includes a sentence from an unrelated paper that contains an algebraic equation. Overall, subjects judged the complex abstract—the one with the nonsense mathematics—to be of higher quality.

Researchers are aware that they can advance their careers by writing in a complex way. MIT students developed SCIGen, computer software to randomly select common but complex words and apply grammar rules to produce documents that pass as research papers on computer science. The title of one paper generated by the software was: “Simulating Flip-flop Gates Using Peer-to-peer Methodologies.” At least 120 such computer-generated papers were published in peer-reviewed scientific journals (Labbé and Labbé, 2013; Lott, 2014).

Juster (1972, p. 23) states, “Few people would accept the naïve no-change model even if it were clearly shown to be more accurate.” This supposition was supported by Hogarth’s (2012) description of four key developments in forecasting in which senior academics resisted overwhelming evidence that simple methods provide forecasts that are more accurate than those from complex ones.

Clients might prefer forecasts that support their plans—another reason for the popularity of complexity in forecasting. Developing complex methods that can be used to provide forecasts that support a desired outcome is relatively easy.

Discussion

During our more than two years working on this special issue, we made repeated requests for experimental evidence that complexity improves forecast accuracy under some conditions. With the enormous efforts and expenditures on analytics and other complex big data methods, one would expect some papers to provide evidence in favor of complexity in forecasting. We have not been able to find such papers, despite our efforts to do so.

Summary of evidence from this review

To obtain a rough idea of the effects of complexity on ex ante forecast accuracy we examined the studies described in this article to find estimates for the direction and size of the effects. The assessments are crude because the definitions of simplicity vary across the different studies. We do not, therefore, claim that our estimate of the effect of complexity on forecast accuracy is definitive, nor do we claim that our review was comprehensive. We hope that other researchers will expand on our work, and perhaps find forecasting problems for which complexity produces forecasts that are substantially more accurate than forecasts from sophisticatedly simple methods. Our judgments are provided at simple-forecasting.com, so others can examine the effects of their own judgments as they see fit.

In total we identify 32 papers—journal articles and book chapters—incorporating 97 formal comparisons of the accuracy of forecasts from complex methods with those from simple—but not in all cases sophisticatedly simple—methods. Eighty-four percent of the comparisons found that forecasts from simple methods were more accurate than, or similarly accurate to, those from complex methods. Averaged across the 25 papers that provide quantitative comparisons, the errors of forecasts from complex methods were 27 percent greater than the errors of forecasts from simple methods. Table 1 summarizes comparisons.

Table 1
Summary of evidence on accuracy of forecasts from complex vs. simple methods

<u>Method type</u>	----- Number of Comparisons -----				<u>Error increase vs simple (%)</u>
	<u>Total papers</u>	<u>Total comparisons</u>	<u>Simple better or similar</u>	<u>Effect size</u>	
Judgmental	4	4	4	4	28.2
Extrapolative	17	62	51	12	27.5
Causal	8	23	19	5	25.3
Combined	3	8	7	4	23.9
All method types	32	97	81	25	
Weighted average*					26.7

*Weighted by total papers

Predicting unusual events

Simple evidence-based methods seem well equipped to deal with the problem of predicting unusual events. For example, Nikolopoulos, Litsa, Petropoulos, Bougioukos, and Khanmash (this issue) extend the research on forecasting special events by testing variations of established sophisticatedly simple methods to forecast the take-up of two new government programs. They found that asking diverse experts to propose and discuss analogies can lead to an error reduction of as much as 54 percent compared to using unaided judgment.

The index method is another simple approach that is well suited to forecasting unusual events. By allowing forecasters to include all variables that are known to be important in a model, an index model is more likely to accurately predict an extreme event than is a statistical model estimated from historical data on a subset of variables that does not include all possible combinations of variable values.

Other evidence on simplicity versus complexity in forecasting

An alternative approach to assess the effects of simplicity in forecasting is to assess the simplicity of evidence-based methods that have been shown to produce accurate forecasts. The Forecasting Methods Selection Tree—available online at forecastingprinciples.com—presents such forecasting methods in the form of a decision tree. The Tree is the product of 40 leading experts in various areas of forecasting from different disciplines, and 123 reviewers (Armstrong 2001), along with evidence cited on the forecastingprinciples.com site since 2001.

None of the 22 methods in the Selection Tree is complex by the definition proposed in this paper. In other words, the evidence-based methods that are recommended for use because they have been shown to offer superior forecasting accuracy are all simple.

Evidence to date favors simplicity in forecasting

The evidence on comparative predictive validity along with the evidence related to the effective methods listed in the Forecasting Methods Selection Tree constitute strong arguments for starting simple when forecasting, and adding complexity only if needed. To the best of the authors of this article's knowledge, the need for complexity has not arisen. Perhaps future research will identify benefits from complexity under some conditions. To establish such a claim would, however, require that researchers test their complex methods against sophisticatedly simple evidence-based methods, and publish their findings regardless of the results.

Work on this *JBR* Special Issue on simplicity in forecasting began with the expectation that simple forecasting would help to improve understanding, reduce mistakes, reveal bias, and identify fraud. The authors expected that forecasts from simple methods would also generally tend to be somewhat more accurate, but were concerned that the research literature would not present a true picture of simplicity in forecasting. In particular, researcher bias toward confirmation of their hypotheses—whether intended or not—would likely produce an overabundance of studies finding in favor of complex methods. Moreover, referee bias against simple methods likely has a damping effect on the number of papers that find in favor of simple methods. The results, then, astonish the authors. The results are consistent across the papers in finding in favor of simple methods, to the extent that complex methods bring with them a forecast error penalty of about one-quarter.

The gains from simplicity identified here are consistent with the gains from following the simple guidelines of the Golden Rule of Forecasting (Armstrong, Green, and Graefe, this issue). In that article, violating a typical guideline increased error by about one-third, and each additional violation would cause further losses in accuracy. The substantial gains in accuracy that are possible from sophisticatedly simple forecasting are a tribute to a half-century of research by the evidence-based forecasting community.

Summary and conclusions

The search for evidence for this introduction to the Special Issue ends with the conclusion that forecasting procedures should always be simple enough for forecast users to understand. Complexity beyond the sophisticatedly simple fails to improve accuracy in all but 16 of the 97 comparisons in 32 papers that provide evidence. Complexity increases forecast error by an average of 27 percent across the 25 papers with quantitative comparisons. In addition to accuracy, simple methods can increase understanding, reduce the likelihood of errors, and aid decision-makers.

Remarkably, no matter what type of forecasting method is used, complexity harms accuracy. Complexity increases errors for forecasts from judgmental, extrapolative, and causal methods by an average of more than 25 percent. Complexity when combining forecasts increases errors by nearly 25 percent. Moreover, all of the 22 useful evidence-based methods found by the forecasting principles project are simple.

Given the weight of evidence and the manifest advantages of simplicity, the advocates of complexity in forecasting surely must be the ones to shoulder the burden of proving that their methods and models will provide forecasts that are accurate *relative to those from sophisticatedly simple ones*. They have so far failed to do so.

This introduction to the Special Issue describes how forecast users can evaluate the simplicity of forecasting methods. A simplicity checklist—which should take only minutes to complete—is available at simple-forecasting.com to guide those evaluations. Obtaining independent ratings of simplicity from several objective raters is desirable, but even a single rating would alert the user to harmful complexity.

If you nevertheless use forecasts from complex methods to help you make decisions, expect to be confused about how the forecasts were made and an accuracy penalty of more than one-quarter.

Forecast accuracy in many fields has failed to improve not because of a lack of knowledge about how to forecast better, but due to a preference for complexity.

References

- Armstrong, J. S. (1970). How to avoid exploratory research. *Journal of Advertising Research*, 19(August), 27–30.
- Armstrong, J. S. (1978). Forecasting with econometric methods. *Journal of Business*, 51(4), 549–564.
- Armstrong, J. S. (1980). Unintelligible management research and academic prestige. *Interfaces*, 10(2), 80–86.
- Armstrong, J. S. (1984). Forecasting by extrapolation: Conclusions from 25 years of research [with Comment and Reply]. *Interfaces*, 14(6), 52–66.
- Armstrong, J. S. (1985). *Long-Range Forecasting*. New York: John Wiley.
- Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. New York: Springer.
- Armstrong, J. S. (2010). *Persuasive Advertising*. New York: Palgrave Macmillan.
- Armstrong, J. S. (2012). Illusions in regression analysis. *International Journal of Forecasting*, 28, 689–694.
- Armstrong, J. S., & Collopy, F. (1992). Error methods for generalizing about forecasting methods. *International Journal of Forecasting*, 8, 69–80.
- Armstrong, J. S., Green, K., & Graefe, A. (2015). The golden rule of forecasting. *Journal of Business Research*, [this issue], xxx–yyy.
- Armstrong, J. S., & Overton, T. (1971). Brief vs. comprehensive descriptions in measuring intentions to purchase. *Journal of Marketing Research*, 8, 114–117.
- Armstrong, J. S., & Shapiro, A. C. (1974). Analyzing quantitative models. *Journal of Marketing*, 38(2), 61–66.
- Boylan, J. E., & Syntetos, A. A. (2006). Accuracy and accuracy-implication metrics for intermittent demand. *Foresight: the International Journal of Applied Forecasting*, 4, 39–42.
- Brighton H., & Gigerenzer, G. (2015). The bias bias. *Journal of Business Research*, [this issue], xxx–yyy.
- Brown, R. G. (1956). Exponential smoothing for predicting demand. Presented at the Tenth National Meeting of the *Operations Research Society of America*, San Francisco, November 16.
- Charlesworth, M. J. (1956). Aristotle's razor. *Philosophical Studies*, 6, 105–112.
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5, 559–583.
- Crone, S. F., Hibon, M., & Nikolopoulos, K. (2011). Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction. *International Journal of Forecasting*, 27, 635–660.
- Cumming, G. (2012). *Understanding the New Statistics: Effect sizes. Confidence Intervals and Meta-Analysis*. New York: Routledge.
- Duncan, G., Gorr, W. L., & Szczypula, J. (2001). Forecasting analogous time series. In J. S. Armstrong (Ed.), *Principles of Forecasting*. Norwell, MA: Kluwer Academic Press.
- Eriksson, K. (2012). The nonsense math effect. *Judgment and Decision Making*, 7(6), 746–749.
- Fader, P. F., & Hardie, B. G. S. (2001). Forecasting trial sales of new consumer packaged goods. In J. S. Armstrong (Ed.), *Principles of forecasting: a handbook for researchers and practitioners* (pp. 613–630). New York: Springer.
- Fildes, R., & Petropoulos, F. (2015). An evaluation of simple versus complex selection rules for forecasting many time series. *Journal of Business Research*, [this issue], xxx–yyy.
- Fildes, R., Wei, Y., & Ismail, S. (2011). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *International Journal of Forecasting*, 27, 902–922.

- García-Ferrer, A., de Juan, A., & Poncela, P. (2006). Forecasting traffic accidents using disaggregated data. *International Journal of Forecasting*, 22, 203–222.
- Gardner, E. S. Jr. (2006). Exponential smoothing: The state of the art—Part II. *International Journal of Forecasting*, 22, 637–666.
- Goodwin, P. (2015). When simple alternatives to Bayes formula work well: Reducing the cognitive load when updating probability forecasts. *Journal of Business Research*, [this issue], xxx–yyy.
- Gorr, W. L., Nagin, D., & Szczypula, J. (1994). Comparative study of neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, 10, 17–34.
- Graefe, A. (2015). Improving forecasts using equally weighted predictors. *Journal of Business Research*, [this issue], xxx–yyy.
- Graefe, A., Küchenhoff, H., Stierle, V., & Riedl, B. (2014). Limitations of ensemble Bayesian model averaging for forecasting social science problems. *International Journal of Forecasting* (forthcoming), available at <http://ssrn.com/abstract=2266307> accessed 30th November 2014.
- Green, K. C. (2005). Game theory, simulated interaction, and unaided judgement for forecasting decisions in conflicts: Further evidence. *International Journal of Forecasting*, 21, 463–472.
- Green, K. C., & Armstrong, J. S. (2007). Structured analogies for forecasting. *International Journal of Forecasting*, 23, 365–376.
- Green, K. C., & Armstrong, J. S. (2011). Role thinking: Standing in other people's shoes to forecast decisions in conflicts. *International Journal of Forecasting*, 27, 69–80.
- Hauer, E. (2004). The harm done by tests of significance. *Accident Analysis and Prevention*, 36, 495–500.
- Hogarth, R. M. (2012). When simple is hard to accept. In P. M. Todd, G. Gigerenzer & A. R. Group (Eds.), *Ecological Rationality: Intelligence in the World* (pp. 61–79). Oxford: Oxford University Press.
- Hogarth, R. M., & Soyer, E. (2015). Communicating forecasts: The simplicity of simulated experience. *Journal of Business Research*, [this issue], xxx–yyy.
- Huddleston, S. H., Porter, J. H., & Brown, D. E. (2015). Improving forecasts for noisy geographic time series. *Journal of Business Research*, [this issue], xxx–yyy.
- Jensen, U. (2001). The simplicity of an earnings frontier. In A. Zellner, H. A. Keuzenkamp, & M. McAleer (Eds.), *Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple* (pp. 277–291). Cambridge: Cambridge University Press.
- Juster, F. T. (1972). An evaluation of the recent record in short-term forecasting. *Business Economics*, 7(3), 22–26.
- Labbé, C., & Labbé, D. (2013). Duplicate and fake publications in the scientific literature: How many SCIdgen papers in computer science? *Scientometrics*, 94, 379–396.
- Lees, G., & Wright, M. (2004). The effect of concept formulation on concept test scores. *Journal of Product Innovation Management*, 21, 389–400.
- Lesser, C. E. V. (1968). A survey of econometrics. *Journal of the Royal Statistical Society: Series A*, 131, 530–566.
- Lott, M. (2014). Over 100 published science journal articles just gibberish. *FoxNews.com*, March 01.
- Lyon, A., Wintle, B. C., & Burgman, M. (2015). Collective wisdom: methods of confidence interval aggregation. *Journal of Business Research*, [this issue], xxx–yyy.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1, 111–153.

- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. F. (1993). The M2-Competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9, 5–22.
- Makridakis, S., & Hibon, M. (2000). The M-3 competition: results, conclusions and implications. *International Journal of Forecasting*, 16, 451–476.
- Malkiel, B. G. (2012). *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*. (10th ed.). New York: W.W. Norton.
- McCloskey, D. N., & Ziliak, S. T. (1996). The standard error of regressions. *Journal of Economic Literature*, 34, 97–114.
- Meade, N., & Islam, T. (2001). Forecasting the diffusion of innovations: Implications for time-series extrapolation. In J. S. Armstrong (Ed.), *Principles of Forecasting*. Boston: Kluwer Academic Publishers.
- Naftulin, D. H., Ware, J. E., Jr., & Donnelly, F. A. (1973). The Doctor Fox lecture: A Paradigm of educational seduction. *Journal of Medical Education*, 48, 630–635.
- Nikolopoulos, K., Goodwin, P., Patelis, A., & Assimakopoulos, V. (2007). Forecasting with cue information: a comparison of multiple regression with alternative forecasting approaches. *European Journal of Operational Research*, 180, 354–368.
- Nikolopoulos, K., Litsa, A., Petropoulos, F., Bougioukos, V., & Khanmash, M. (2015). Relative performance of methods for forecasting Special Events. *Journal of Business Research*, [this issue], xxx–yyy.
- OED Online (2014). “simplicity, n.”. Oxford University Press. <http://www.oed.com/view/Entry/179975?redirectedFrom=simplicity&> accessed 8th October 2014.
- Peach, J. T., & Webb, J. L. (1983). Randomly specified macroeconomic models: some implications for model selection. *Journal of Economic Issues*, 17, 697–720.
- Schmidt, F. L. (1971). The relative efficiency of regression and simple unit predictor weights in applied differential psychology. *Educational and Psychological Measurement*, 31, 699–714.
- Schnaars, S. P. (1984). Situational factors affecting forecast accuracy. *Journal of Marketing Research*, 21, 290–297.
- Schnaars, S. P., & Bavuso, R. J. (1986). Extrapolation models on very short-term forecasts. *Journal of Business Research*, 14(1), 27–36.
- Shiskin, J. (1965). *The X-11 variant of the Census Method II seasonal adjustment program*. Washington, D.C.: U.S. Bureau of the Census.
- Smith, S. K. (1997). Further thoughts on simplicity and complexity in population projection models. *International Journal of Forecasting*, 13, 557–565.
- Soyer, E., & Hogarth, R. (2012). Illusion of predictability: How regressions statistics mislead experts. *International Journal of Forecasting*, 28, 695–711.
- Syntetos, A. A., Babai, M. Z., & Gardner, E. S. (2015). Forecasting intermittent inventory demands: Simple parametric methods vs. bootstrapping. *Journal of Business Research*, [this issue], xxx–yyy.
- Tessier, T. H., & Armstrong, J. S. (2015). Decomposition of time-series by current level and change. *Journal of Business Research*, [this issue], xxx–yyy.
- Teunter, R. H., & Duncan, L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society*, 60, 321–329.
- Willemain, T. R., Smart, C. N., & Schwarz, H. F. (2004). A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting*, 20(3), 375–387.
- Woike, J. K., Hoffrage, U., Petty, J. S. (2015). Picking profitable investments: The success of equal weighting in simulated venture capitalist decision making. *Journal of Business Research*, [this issue], xxx–yyy.

- Wright, M. J., & Stern, P. (2015). Forecasting new product trial with analogous series. *Journal of Business Research*, [this issue], xxx–yyy.
- Zellner, A. (2001). Keep it sophisticatedly simple. In A. Zellner, H. A. Keuzenkamp, & M. McAleer (Eds.) *Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple* (pp. 242–262). Cambridge: Cambridge University Press.
- Zellner, A., Keuzenkamp, H. A., & McAleer, M. (Eds.) (2001). *Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple*. Cambridge: Cambridge University Press.
- Ziliak, S. T., & McCloskey, D. N. (2004). Size matters: the standard error of regressions in the American Economic Review. *The Journal of Socio-Economics*, 33, 527–546.
- Ziliak, S. T., & McCloskey, D. N. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor: The University of Michigan Press.

Table 1
Summary of evidence on accuracy of forecasts from complex vs. simple methods

<u>Method type</u>	----- Number of Comparisons -----				<u>Error increase vs simple (%)</u>
	<u>Total papers</u>	<u>Total comparisons</u>	<u>Simple better or similar</u>	<u>Effect size</u>	
Judgmental	4	4	4	4	28.2
Extrapolative	17	62	51	12	27.5
Causal	8	23	19	5	25.3
Combined	3	8	7	4	23.9
All method types	32	97	81	25	
Weighted average*					26.7

*Weighted by total papers