Improving forecasts using equally weighted predictors

Andreas Graefe
LMU Research Fellow
Center for Advanced Studies
Department of Communication Science and Media Research
LMU Munich, Germany
a.graefe@lmu.de

Abstract. The usual procedure for developing linear models to predict any kind of target variable is to identify a subset of most important predictors and to estimate weights that provide the best possible solution for a given sample. The resulting "optimally" weighted linear composite is then used when predicting new data. This approach is useful in situations with large and reliable datasets and few predictor variables. However, a large body of analytical and empirical evidence since the 1970s shows that the weighting of variables is of little, if any, value in situations with small and noisy datasets and a large number of predictor variables. In such situations, including all relevant variables is more important than their weighting. These findings have yet to impact many fields. This study uses data from nine established U.S. election-forecasting models whose forecasts are regularly published in academic journals to demonstrate the value of weighting all predictors equally and including all relevant variables in the model. Across the ten elections from 1976 to 2012, equally weighted predictors reduced the forecast error of the original regression models on average by four percent. An equal-weights model that includes all variables provided well-calibrated forecasts that reduced the error of the most accurate regression model by 29% percent.

Keywords: equal weights, index method, econometric models, presidential election forecasting

Acknowledgements: J. Scott Armstrong and Alfred Cuzán provided helpful comments.

1. Introduction

People and organizations commonly make decisions by combining information from multiple inputs. For example, one usually weighs the pros and cons before deciding on whether or not to launch a marketing campaign, which new product to develop, or where to open a branch office. Almost 250 years ago, Benjamin Franklin suggested an approach for how to solve such problems. Franklin's friend Joseph Priestley asked for advice on whether or not to accept a job offer that would have involved moving with his family from Leeds to Wiltshire. In his response letter, written on September 19, 1772, Franklin avoided advising Priestley on *what* to decide. Instead, he proposed a method for *how* to decide. Franklin's recommendation was to list all important variables, decide which decision is favored by each variable, weight each variable by importance, and then add up the variable scores to see which decision is ultimately favored. Franklin' labeled this approach "Moral Algebra, or Method of deciding doubtful Matters" (Sparks, 1844, p. 20). About half a century later, Franklin's method had another famous proponent. In 1838, Charles Darwin used it to help him answer a question of utmost importance: whether or not to get married (Darwin, Burkhardt, & Smith, 1986).

Franklin's Moral Algebra gave way to multiple regression analysis, which has become popular for solving many kinds of problems in various fields. Multiple regression analysis produces variable weights that yield the "optimal" (in terms of least squares) solution for a given data set. The estimated regression coefficients are then commonly used to weight the composite when predicting new (out-of-sample) data. The problem with this data fitting approach is that it does not necessarily yield accurate forecasts. A large body of empirical and theoretical evidence since the 1970s shows that regression weights often provide *less* accurate out-of-sample forecasts than simply assigning equal weights to each variable in a linear model (Dawes, 1979; Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975). These results have yet to impact many fields, including business research. Researchers rarely evaluate the quality of their models by predicting holdout data and most JBR submissions report the model fit as the only indication of a good model (Woodside, 2013).

I review the literature on the relative predictive performance of equal and regression weights and provide new evidence from the field of U.S. presidential election forecasting, a field that is dominated by the application of multiple regression analysis. The results conform to prior research, showing that equal weights perform at least as well as regression weights when forecasting new data. In addition, I show that including all relevant variables in an equal-weights model yields large gains in accuracy.

2. Equal and regression weights in linear models

This section reviews prior research on the relative performance of equal and regression weights and discusses the conditions under which either approach is expected to work best.

2.1 Multiple regression models

As mentioned above, multiple regression analysis is the dominant method to develop forecasting models in many fields. Once theory is used to select the *k* relevant predictor variables, multiple regression analysis estimates their relative impact on the target criterion. The general equation of the multiple regression model reads as:

$$y = a + \sum_{i=1}^{k} b_i x_i + e \qquad (1)$$

The estimated constant *a* and the *k* "optimal" (in terms of minimized squared error) regression coefficients b_i are then used when predicting new data.

2.2 Equal-weights models

An alternative to using multiple regression is to assign equal weights to each variable. That is, one also relies on theory to select the variables. However, one does not let the data decide about the variables' weights. Instead, one uses prior knowledge to assess the directional effects of the variables and then transforms all variables so that they positively correlate with the target variable. In the final step, the values of all variables are added up to calculate the single predictor variable in a simple linear regression model, hereafter, the equal-weights model:

$$y = d + g\sum_{i=1}^{k} x_i + v \qquad (2)$$

where *d* is the estimated constant, *g* is the estimated coefficient of the predictor variable, and *v* is the error term.

2.3 Differences between multiple regression and equal-weights models

The multiple regression and the equal-weights model differ in the number of parameters to be estimated. The multiple regression model estimates k+1 parameters: the constant *a* and each variable's coefficient b_i. The equal-weights model is a special case of the multiple regression model with all b_i's = *g*. That is, the equal-weight method only needs to estimate two parameters (*d* and *g*).

The number of estimated parameters is crucial to a model's predictive performance, since their estimation inevitably creates error. While adding more variables will generally improve a model's fit to existing data, the danger of overfitting increases. Overfitted models tend to exaggerate minor fluctuations in the data by interpreting noise as information. As a result, the models' performance for predicting new data *de*creases. The equal-weights model uses as few degrees of freedom as possible and thus minimizes estimation error. The relative performance of multiple regression and equal-weights models for the same data then depends on the accuracy of the estimated coefficients. See Einhorn and Hogarth (1975) for a more detailed discussion.

2.4 Empirical evidence on the relative performance of multiple regression and equal-weights models

Starting at least as early as Schmidt (1971), a number of studies have tested the relative predictive accuracy of equal and regression weights when applied to the same data. Many of these studies analyzed unit weights, which are a special case of equal weights in which each variable is assigned a value of plus or minus one.

An early review of the literature finds multiple regression to be slightly more accurate than equal weights in three studies but less accurate in five (Armstrong, 1985, p. 208). Since then, evidence has accumulated. Czerlinski, Gigerenzer, and Goldstein (1999) test the predictive performance of regression and equal weights for twenty real-world problems in areas such as psychology, economics, biology, and medicine. Most of these tasks were collected from statistics textbooks where they were used to demonstrate the application of multiple regression analysis. Ironically, equal weights provided more accurate predictions than multiple regression. Cuzán and Bundrick (2009) analyze the relative performance of equal and regression weights for forecasting U.S. presidential elections. The authors find that equal-weights versions of the Fair (2009) model and of two variations of the fiscal model (Cuzán & Heggen, 1984) outperformed two of the three regression models – and did equally well as the third – when making out-of-sample predictions.

Such findings have led researchers to conclude that the weighting of variables is secondary for the accuracy of forecasts. Once the relevant variables are included and their directional impact on the criterion is specified, the magnitudes of effects are not very important (Armstrong, 1985, p. 210; Dawes, 1979). As Dawes and Corrigan (1974, p. 105) put it in their seminal work on that topic: "The whole trick is to decide which variables to look at and then to know how to add."

2.5 Conditions for the relative performance of multiple regression and equal-weights models

The relative performance of equal and regression weights depends on the conditions of the forecasting problem. Analytical solutions to the problem derived several conditions for when equal weights can outperform regression weights when predicting new data (Davis-Stober, Dana, & Budescu, 2010; Einhorn & Hogarth, 1975). These conditions are common for many problems in the social sciences. In general, the relative performance of equal weights increases if

1. the regression model fits the data poorly
   (i.e., the multiple correlation coefficient R^2 is low),
2. the ratio of observations per predictor variable is low
   (i.e., in situations with small samples and a large number of predictor variables),
3. predictor variables are highly correlated, and
4. there is measurement error in the predictor variables.

Empirical studies yield similar conclusions. Dana and Dawes (2004) analyze the relative predictive performance of regression and equal weights for five real non-experimental social science datasets and a large number of synthetic datasets. They find that regression weights do not yield more accurate forecasts than equal weights unless sample size is larger than *one hundred observations per predictor*. Only in cases in which prediction error was likely to be very small (adjusted R^2>.9), the authors found regression to outperform equal weights in samples with five observations per predictor.

3. Models for forecasting U.S. presidential elections

The development of quantitative models to predict the outcome of elections is a well-established sub-discipline of political science. Since the late 1970s, scholars have developed various versions of election forecasting models. Table 1 shows the specifications of nine models, including the variables used, their first election forecasted, the sample period, and the model fit. The figures are based on data up to – but not including – the 2012 election. That is, the model specifications show the situation that the forecasters faced prior to the 2012 election.

Eight of the nine models are described in *PS: Political Science & Politics* 45(4). The latest specification of the Fair model is described in Fair (2009). Also, note that the specification of the Abramowitz model differs from the author's description in *PS: Political Science & Politics* 45(4). In his article, Abramowitz proposed a revised model with one additional variable. However, at the SPSA 2013 meeting in Orlando, Abramowitz indicated that he would likely return to his old model in the future. Therefore, the present analysis stays with the established "time-for-change" model.

Each of these models is estimated using multiple regression analysis, with the incumbent popular two-party vote as the dependent variable and two or more independent variables derived from theory. For example, it is well known that elections can be viewed as referenda about the government's performance or, more narrowly defined, its ability to handle the economy. That is, voters reward the government for good performance and punish the incumbent party otherwise. Most models incorporate this information by using one or more economic variables (e.g., GDP growth or job creation) to measure economic performance. Other popular measures are presidential popularity, which is commonly seen as a proxy variable for measuring the incumbent's overall performance, and the time the incumbent party has held the White House. The models are then used to test theories of voting, to estimate the relative effects of specific variables on the aggregate vote, and, of course, to forecast the election outcome.

The conditions for forecasting U.S. presidential elections suggest that the equal-weights method should perform well compared to multiple regression. Although most of the nine models listed in Table 1 fit the data fairly well, the number of observations per predictor variable is low, the predictors are likely correlated, and forecasters have to deal with measurement errors in the predictor variables.

Model fit. Models for forecasting U.S. presidential elections are able to explain much of the variance in the two-party popular vote shares. Table 1 shows the fit of the nine models, estimated using data up to and including the 2012 election. Seven of the nine models explain more than 80% of the variance; one model (Cu) achieves an adjusted R^2 above .9.

Ratio of observations per predictor variable. Seven of the nine models listed in Table 1 use only data post-World War II. This means that these models were limited to around fifteen observations when estimating the vote equation to predict the 2012 election results. The two exceptions are the models by Fair and Cuzán, which start collecting data in 1916, and thus drew on a sample of twenty-four observations when calculating the forecast of the 2012 election. The number of predictor variables differs across models. While four models are based on two variables, the Fair model uses seven variables. Thus, when calculating forecasts of the 2012 election, the ratio of observations per predictor ranged from 3.4 (F) to 8.0 (C). These ratios are far below what Dana and Dawes (2004) recommended when using multiple regression. In addition, these ratios were of course lower for forecasts of earlier elections.

Correlation among predictors. In most real-world forecasting problems, predictor variables are likely correlated. This also holds for election forecasting. An example is the combined use of economic indicators and public opinion polls (e.g., presidential popularity) as predictor variables in the same model. Since presidential popularity is expected to serve as a proxy for incumbent performance, the measure likely also captures the public's perceptions of how the president is handling the economy. Prior research supports this. Ostrom and Simon (1985) find that presidential popularity is a function of both economic and non-economic factors. Similarly, Lewis-Beck and Rice (1992, p. 46) show that GNP growth is correlated with incumbent performance ($r = .48$). Five of the nine models listed in Table 1 use both economic indicators and public opinion polls (A, C, EW, Ho, and LT).

Measurement error in independent variables. As shown in Table 1, economic indicators and public opinion polls are major predictors in election forecasting models; both measures are subject to measurement error.

First, the state of the economy is difficult to measure. Often, there are substantial differences between initial and revised estimates of economic figures. For example, on January 30, 2009, the *Bureau of Economic Analysis* at the *U.S. Department of Commerce* initially estimated a real GDP decrease of 3.8 percent for the fourth quarter of 2008. One month later, the figure was revised to 6.2 percent, and, at the time of writing, the latest estimate showed a decrease of 8.9 percent. Revisions of this size are not exceptional. Runkle (1998) analyzes deviations between initial and revised estimates of quarterly GDP growth from 1961 to 1996. Revisions were common. There were upward revisions by as much as 7.5 percentage points and downward revisions by as much as 6.2 percentage points. Such measurement errors are even more critical when different estimates are used for building a model and calculating the forecast. For example, forecasters commonly use revised economic figures to estimate the model. However, when making the forecast shortly before the election, the forecasters have to draw on the initial estimates, since the revised figures are not yet available.

Second, polls conducted by reputable survey organizations at about the same time often reveal considerable variation in results. Errors caused by sampling problems, non-responses, inaccurate measurement, and faulty processing diminish the accuracy of polls and the quality of surveys more generally (Erikson & Wlezien, 1999). Such measurement errors can have a large impact on the validity of the estimated model coefficients and thus on the accuracy of forecasting models.

4. Evidence on the accuracy of multiple regression and equal-weights models in forecasting U.S. presidential elections

The following analysis extends prior work by Cuzán and Bundrick (2009) and tests the predictive performance of equal and regression weights for the models listed in Table 1.

4.1 Method

All data and calculations are available at: tinyurl.com/equalweights.

4.1.1 Models and data

The present study analyses forecasts from the nine models listed in Table 1. Data for six models (A, Ca, EW, F, Hi, LT) were obtained from Montgomery, Hollenbach, and Ward (2012) and enhanced with the variable values from the 2012 election. The data for the model by Lockerbie were derived from Lockerbie (2012). Thanks to Alfred Cuzán and Thomas Holbrook who shared their data.

For the purpose of this study it was necessary to perform some transformations on the original data. Without any loss of generality, the data were transformed in standardized (z-scores) format such that each predictor correlates positively with the dependent variable. The dependent variable was the two-party popular vote received by the candidate of the incumbent party.

4.1.2 Forecast calculations

All forecasts analyzed in the present study can be considered *pseudo ex ante*, calculated as one-election-ahead predictions. That is, only data that *would* have been available at the time of the particular election being forecast was used to estimate the model. For example, to calculate a forecast for the 2004 election, only data up to the 2000 election was used to estimate the model. To calculate a forecast of the 1984 election, only data up to 1980 was used, and so on.

The term "pseudo" reveals that these forecasts cannot be considered truly ex ante. The reason is that all calculations are based on the models' specifications that were used for predicting the 2012 elections. In reality, however, the 2012 versions were often quite different from the original specifications that were used to predict a particular election. Most models have been revised at least once since their first publication, usually as a reaction to poor performance in forecasting the previous election. Such revisions usually improve the fit of the regression model to historical data. As a result, the pseudo ex ante forecasts tend to be more accurate than what one would have obtained with the original model specifications that were used in the actual elections. The only exception are the forecasts of the 2012 election, which can be considered as "truly" ex ante, since they are only based on information that was actually available at the time of making the forecast. The interested reader can track how most of the model specifications have changed over time by referring to the forecasters' manuscripts, which were published prior to each election since 1992 in special symposiums of *Political Methodologist* 5(2), *American Politics Research* 24(4) and *PS: Political Science and Politics* 34(1), 37(4), 41(4), and 45(4).

The *multiple regression model* forecasts represent the forecasts of the 2012 model specifications. That is, multiple regression analysis was used to regress the incumbent party's popular two-party vote share on the set of independent variables included in each model (cf. equation 1). The *equal-weights model forecasts* represent the forecasts of the equal-weights variant of each model. This was done by summing up the scores of the predictor variables incorporated in each model. The resulting equal-weights score was then used as the single predictor variable in a simple linear regression model (cf. equation 2).

4.1.3 Forecast horizon and error measure

Forecast accuracy was analyzed across the ten U.S. presidential elections from 1976 to 2012. The absolute error was used to measure the absolute deviation of the forecast from the actual election result. The error reduction was used to compare the relative performance of forecasts based on equal and regression weights. The error reduction is simply the difference between the absolute errors of the multiple regression and the equal-weights forecasts. Negative values mean that regression weights provided more accurate forecasts than equal weights. Positive values mean that equal weights outperformed regression weights.

4.2 Results

Table 2 shows the mean absolute error (MAE) of the multiple regression and the equal-weights variants of each model, as well as their relative accuracy, measured as the error reduction in percentage points (and in percent).

Across all ten elections from 1976 to 2012, there were little differences in the relative accuracy of equal and regression weights. The equal-weights models were more accurate than the multiple regression models in six cases (A, Cu, EW, F, Ho, L) and less accurate in three cases (Ca, LT, Hi). On average across the nine models, the error of the equal-weights models was 0.1 percentage points lower than the corresponding error of the regression models.

However, the multiple regression models have an advantage in this comparison. None of the nine models was around in 1976, and some were developed not until the late 1990s (e.g., Cu, Hi, L, LT). Therefore, forecasts for elections held *before* a model was developed cannot even be considered pseudo ex ante. The reason is that information from subsequent elections was used to *select* the variables when building the model. In order to somewhat account for this problem, Figure 1 shows the MAE of the forecasts from the nine multiple regression models and the corresponding equal-weights variants, both for elections before each model was first developed and since its first application. The results suggest that multiple regression models benefit from data fitting. When "predicting" elections that were held before the model was created, regression weights were slightly more accurate than equal weights. However, for elections held since the models were first created, the results are reversed. The accuracy loss of regression models when predicting data that were unknown to the forecasters at the time they decided about the model specifications is a sign for overfitting (illustrated by the steep slope of the black line in Figure 1). In comparison, the slope for the equal-weights models (grey line) is flatter. Since the equal-weights method only estimates two parameters, the risk of overfitting is smaller. Equal-weights models are more robust and do not suffer from large losses in accuracy when predicting new data.

4.3 Discussion

Equal weight versions of established multiple regression models for forecasting U.S. presidential elections were found to predict as well as – or better than – the original models. Thereby, it is important to emphasize that these results likely *under*estimate the gains in accuracy that can be achieved by using equal instead of regression weights, since the analysis was based on the 2012 specifications of the models (see Section 4.1.2). Thus, the results should be regarded as a low boundary. The actual gains that one would have obtained by using equal instead of regression weights at the time the forecast was made are likely to be higher than the gains reported in Table 2.

These results may surprise given that all of the models that are established in the political science community (i.e., the models that are regularly published in special issues of political science journals) are regression models.[1] However, the results conform to a large body of research since the 1970s that provides empirical and analytical evidence in favor of equal weights when making out-of-sample forecasts for social science problems. This work concluded that, once the relevant variables are identified, the issue of how to weight variables is not critical for forecast accuracy. Much of this research was done years before researchers developed the first U.S. presidential election forecasting models. Since then, evidence showing that equal weights often outperform regression for social science problems has accumulated, also for the domain of election forecasting (Cuzán & Bundrick, 2009); the present study adds more evidence.

Unfortunately, these findings had little impact on election forecasting thus far. Although researchers increasingly demonstrate the usefulness and accuracy of equal-weights models for election forecasting (Armstrong & Graefe, 2011; Graefe & Armstrong, 2013; Lichtman, 2008), these findings are rarely published in political science journals.

The present study does not mean to suggest that regression cannot be useful for forecasting. Regression analysis is an important forecasting method and there is much evidence that it can provide accurate predictions if used under appropriate conditions (Allen & Fildes, 2001; Armstrong, 1985). The method is particularly useful in situations with large

[1] This is not to say that there are no equal-weights models for forecasting U.S. presidential elections. The most popular equal-weights model is Lichtman's "Keys to the White House", which is based on

reliable datasets, few variables that are based on well-established causal relationships with the criterion and do not highly correlate with each other, and when the expected changes are large and predictable (Armstrong, 2012).

The problem is that these conditions are rarely met when predicting social science problems. Rather, the conditions often favor equally weighted predictors. Given its demonstrated accuracy and obvious simplicity, it is surprising that the equal-weights method has been widely overlooked, not to say ignored.

A common objection to the equal-weights method is that the use of equal weights is considered unscientific or atheoretical. The likely reason for this objection is that the outputs of equal-weights models do not conform to what users of regression analysis expect to see. In particular, the equal-weights method does not estimate effect sizes and therefore cannot provide answers to questions such as whether a variable has a statistically significant impact, how large this effect is, or whether one variable is more important than another. Given that users of regression analysis often argue that the main purpose of their model is not to forecast but to test theory and to estimate the size of effects, this appears to be a major limitation of the equal-weights method. But should we have faith in the validity of effect sizes that are little or no more accurate than equal weights when predicting new data? And is not, after all, the best test of a model's validity its predictive accuracy?

Furthermore, users of regression analysis commonly assume that they can control for the relative impact of variables that they put in the equation. However, this assumption only holds for experimental data. For non-experimental data, variables often correlate with (combinations of) other variables; a problem that gets worse when the number of variables increases. Armstrong (2012) refers to this as the "illusion of control" in regression analysis. He recommends that one should not estimate effect sizes from non-experimental data. Instead, one should rely on experiments to estimate effect sizes and then incorporate this information in the model.

Finally, it is a misperception that the equal-weights method prevents analysts from using and testing theory. Rather, analysts *need* to draw on theory and prior knowledge when they select and code variables. To some extent, the equal-weights method is more useful to test theory than multiple regression, since it can include an unlimited number of variables and does not let the data decide about the variables' (directional) effect on the target criterion. The possibility to include all relevant variables in a model is thus one of the major benefits of the equal-weights method, and will be discussed in the following.

5. Including all available information

The above analyses are similar to prior research on the relative performance of equal and regression weights in that they compare both methods by *using the same data*. However, this comparison conceals a major advantage of the equal-weights method. While multiple regression analysis is limited in the number of variables that can be included in a model (cf. Section 2.5), the number of parameters that need to be estimated in an equal-weights model is independent from the number of predictor variables (cf. equation 2). That is, with the equal-weights method one can follow Benjamin Franklin's advice and use *all* relevant variables. The use of all relevant variables is also one of the guidelines in the Golden Rule of Forecasting Checklist (Armstrong, Green, & Graefe, in press).

5.1 Method

To test this approach while holding the data set constant, the independent variables were restricted to those that were used by the nine models analyzed above. As shown in Table 1, the nine models use a total of 30 variables. However, two models (F and Cu) use three identical variables, which reduces the number of unique variables to 27. The sum of these 27 variable values was used as the single predictor variable in a simple linear regression model, hereafter referred to as the *index model*, with the incumbent's popular two-party vote share as the dependent variable. The index model was estimated based on data starting in 1952, since this is the first election for which data on all variables were available. Pseudo ex ante forecasts were again calculated as one-election-ahead predictions.

5.2 Results

As shown in Table 3, the index model provided highly accurate forecasts. The index model's mean absolute error across the ten elections from 1976 to 2012 was 1.3 percentage points. Compared to the individual models, error reductions ranged from 0.5 to 2.7 percentage points. That is, the index model reduced the error of the most accurate individual model (Ca) by 29%; compared to the least accurate model (L), error reduction reached 67%. Compared to the typical model, the index model reduced the error by 48%.

Figure 2 shows the calibration of the forecasts from the index model and eight individual models.[2] The marker shows each model's point forecast, the vertical lines show their 95% prediction intervals, and the dashed horizontal line shows the actual election result.[3] The index model is well calibrated. For nine of the ten elections, the election result falls within the 95% prediction interval. The exception is the 2000 election, in which the index model over-predicted Gore's vote share by a small margin. In addition, the prediction intervals provided by the index model are narrow, which is an important quality criterion for a forecast model. Across the ten elections, the average prediction interval is little larger than five percentage points, which is the lowest value of all models. That is, if the index model

[2] The data to calculate the prediction intervals for the Hibbs model were not available.
[3] The 95% prediction intervals were calculated as twice the standard error.

predicts the incumbent to gain 53 percent of the vote, there is a 95% chance that the actual election result will be between 50.5 and 55.5 percent. In comparison, the prediction interval for the second most accurate model (Ca) spans a range of almost eight percentage points, which makes its forecasts more vague and thus less valuable. Prediction intervals for other models are even wider.

5.3 Discussion

The index model reduced the error of the most accurate individual model by 29% and cut the error of the typical model nearly in half. In addition, the index model forecasts were well calibrated. These large gains in accuracy were achieved by aggregating all information included in the individual models in a single index variable.

These results are consistent with prior research. Researchers have concluded from comparative studies that having all relevant variables in a model is more important than the "optimal" weighting of a set of variables (Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975). The equal-weights method enables analysts to include an unlimited number of variables in a model. This is the most important feature of this method when dealing with situations in which there are many important variables; a situation that is common for social science problems.

However, one might object to the equal-weights method *because* it can incorporate a large number of variables. Parsimony is commonly regarded an important quality criterion of a forecasting model (Lewis-Beck, 2005). However, parsimony is only crucial for methods that need to estimate many parameters and thus bear the risk of overfitting, such as regression analysis. In comparison, the number of variables is no concern for equal-weights models, since the index method does not estimate multiple variable weights (cf. equation 2). In fact, as demonstrated above, it is one of the major benefits of the equal-weights method to be able to include all relevant knowledge.[4]

Finally, one might be concerned about unequal distribution of variables. For example, the index model incorporates thirteen economic variables, six political variables, and eight measures of public opinion. Thus, one might think that economic variables are overrepresented, whereas public opinion polls are underrepresented. Here, it helps to think of the index model as an index of indexes. For example, before calculating the single index variable, one could sum up all economic variables in an economic index, all poll variables in a poll index, all political variables in a political index, and so on. How one aggregates the variable values does not matter; mathematically the results are the same.

[4] Note that the index model incorporates aspects of retrospective and prospective voting, the influence of incumbency, the time-for-change effect, and military losses.

The performance of the index model is heartening and the findings are relevant for many applications, also in the field of business research. The index method is particularly useful for situations with many important variables and if there is prior knowledge about the directional impact of the variables on the target criterion. Prior knowledge can be obtained from empirical evidence, expert knowledge, or, ideally, experimental studies (Graefe & Armstrong, 2011). For example, one study develops an index model to predict the effectiveness of advertisements from 195 evidence-based persuasion principles. When using this model, advertising novices provided more accurate predictions of ad effectiveness than experts' unaided judgment (Armstrong, Rui, Graefe, Green, & House, 2012).

6. Concluding remarks

Benjamin Franklin's advice was to identify *all* variables that are considered important for the problem at hand. And, although he suggested weighting variables by importance, his advice was pragmatic and simple: use intuition. In contrast to many contemporary researchers, Franklin seemed to be little concerned about *how* to estimate optimal variable weights. (With all due respect for Franklin's "Moral Algebra", however, the use of intuition is likely to be harmful to the accuracy of results. The reason is that such an informal weighting approach allows people to assign weights in a way that suits their biases (Graefe, Armstrong, Jones Jr., & Cuzán, 2013).)

Time proved Franklin right. A large body of analytical and empirical evidence from various fields found that the weighting of variables in linear models is uncritical for forecast accuracy; what is most important is to include all relevant variables. Therefore, a good rule of thumb for weighting composites in linear models is to keep things simple and to use equal weights, although differential weights may be useful under certain conditions.

7. References

Allen, P. Geoffrey, & Fildes, Robert. (2001). Econometric forecasting. In J. S. Armstrong (Ed.), *Forecasting Principles: A Handbook for Researchers and Practitioners* (pp. 301-362). New York: Springer.

Armstrong, J. Scott. (1985). *Long-range Forecasting: From Crystal Ball to Computer*. New York: Wiley.

Armstrong, J. Scott. (2012). Illusions in regression analysis. *International Journal of Forecasting, 28*(3), 689-694.

Armstrong, J. Scott, & Graefe, Andreas. (2011). Predicting elections from biographical information about candidates: A test of the index method. *Journal of Business Research, 64*(7), 699-706. doi: 10.1016/j.jbusres.2010.08.005

Armstrong, J. Scott, Green, Kesten C., & Graefe, Andreas. (in press). Golden Rule of Forecasting. *Journal of Business Research*.

Armstrong, J. Scott, Rui, Du, Graefe, Andreas, Green, Kesten C., & House, Alexandra. (2012). Predictive validity of evidence-based advertising principles. *Working paper*.
Cuzán, Alfred G., & Bundrick, Charles M. (2009). Predicting Presidential Elections with Equally Weighted Regressors in Fair's Equation and the Fiscal Model. *Political Analysis, 17*(3), 333-340.
Cuzán, Alfred G., & Heggen, Richard J. (1984). A fiscal model of presidential elections in the United States, 1880-1980. *Presidential Studies Quarterly, 14*(1), 98-108.
Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer & P. M. Todd (Eds.), *Simple heuristics that make us smart* (pp. 97-118): Oxford University Press.
Dana, Jason, & Dawes, Robyn M. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics, 29*(3), 317-331.
Darwin, Charles, Burkhardt, Frederick, & Smith, Sydney. (1986). *The correspondence of Charles Darwin: Volume 2, 1837-1843*. Cambridge: Cambridge University Press.
Davis-Stober, Clintin P., Dana, Jason, & Budescu, David V. (2010). A constrained linear estimator for multiple regression. *Psychometrika, 75*(3), 521-541.
Dawes, Robyn M. (1979). The robust beauty of improper linear models in decision making. *American psychologist, 34*(7), 571-582.
Dawes, Robyn M., & Corrigan, Bernard. (1974). Linear models in decision making. *Psychological Bulletin, 81*(2), 95-106.
Einhorn, Hillel J., & Hogarth, Robin M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance, 13*(2), 171-192. doi: 10.1016/0030-5073(75)90044-6
Erikson, Robert S., & Wlezien, Christopher. (1999). Presidential polls as a time series: the case of 1996. *Public opinion quarterly, 63*(2), 163-177.
Fair, Ray C. (2009). Presidential and congressional vote-share equations. *American Journal of Political Science, 53*(1), 55-72.
Graefe, Andreas, & Armstrong, J. Scott. (2011). Conditions under which index models are useful: Reply to bio-index commentaries. *Journal of Business Research, 64*(7), 693-695.
Graefe, Andreas, & Armstrong, J. Scott. (2013). Forecasting elections from voters' perceptions of candidates' ability to handle issues. *Journal of Behavioral Decision Making, 26*(3), 295-303.
Graefe, Andreas, Armstrong, J. Scott, Jones Jr., Randall J., & Cuzán, Alfred G. (2013). Combining forecasts: An application to elections. *Forthcoming in the International Journal of Forecasting*, ssrn.com/abstract=1902850.
Lewis-Beck, Michael S. (2005). Election forecasting: principles and practice. *The British Journal of Politics & International Relations, 7*(2), 145-164.
Lewis-Beck, Michael S., & Rice, Tom W. (1992). *Forecasting Elections*. Washington, DC: Congressional Quarterly Press.
Lichtman, Allan J. (2008). The keys to the white house: An index forecast for 2008. *International Journal of Forecasting, 24*(2), 301-309.
Lockerbie, Brad. (2012). Economic expectations and election outcomes: The Presidency and the House in 2012. *PS: Political Science & Politics, 45*(4), 644-647.

Montgomery, Jacob M., Hollenbach, Florian, & Ward, Michael D. (2012). Improving predictions using ensemble Bayesian model averaging. *Political Analysis, 20*(3), 271-291.

Ostrom, Charles W. Jr., & Simon, Dennis M. (1985). Promise and performance: A dynamic model of presidential popularity. *American Political Science Review, 79*(2), 334-358.

Runkle, David E. (1998). Revisionist history: how data revisions distort economic policy research. *Federal Reserve Bank of Minneapolis Quarterly Review, 22*(4), 3-12.

Schmidt, Frank L. (1971). The relative efficiency of regression and simple unit predictor weights in applied differential psychology. *Educational and Psychological Measurement, 31*(3), 699-714.

Sparks, Jared. (1844). *The Works of Benjamin Franklin* (Vol. 8). Boston: Charles Tappan Publisher.

Woodside, Arch G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research, 66*(4), 463-472. doi: http://dx.doi.org/10.1016/j.jbusres.2012.12.021