

Dartmouth College

## Dartmouth Digital Commons

---

Dartmouth Scholarship

Faculty Work

---

2015

### Systematic Bias and Nontransparency in US Social Security Administration Forecasts

Konstantin Kashin  
*Harvard University*

Gary King  
*Harvard University*

Samir Soneji  
*Dartmouth College*

Follow this and additional works at: <https://digitalcommons.dartmouth.edu/facoa>



Part of the [Finance Commons](#)

---

#### Dartmouth Digital Commons Citation

Kashin, Konstantin; King, Gary; and Soneji, Samir, "Systematic Bias and Nontransparency in US Social Security Administration Forecasts" (2015). *Dartmouth Scholarship*. 2430.  
<https://digitalcommons.dartmouth.edu/facoa/2430>

This Article is brought to you for free and open access by the Faculty Work at Dartmouth Digital Commons. It has been accepted for inclusion in Dartmouth Scholarship by an authorized administrator of Dartmouth Digital Commons. For more information, please contact [dartmouthdigitalcommons@groups.dartmouth.edu](mailto:dartmouthdigitalcommons@groups.dartmouth.edu).

# Systematic Bias and Nontransparency in US Social Security Administration Forecasts<sup>†</sup>

Konstantin Kashin, Gary King, and Samir Soneji

Since the passage of the Social Security Act of 1935, a central concern of the Board of Trustees has been the demographic and financial forecasts necessary to assess the long-term solvency of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds. These forecasts are used in a variety of ways. For example, they affect decisions about whether the rate of payroll taxes or the amount of benefits should be raised or lowered for 212 million workers and 59 million beneficiaries in 2014, respectively. The methodology rooted in the forecasts is used by the Social Security Administration to evaluate policy proposals put forward by Congress to modify the program. The forecasts are also used as essential inputs in assessing the finances of Medicare and Medicaid and are central to research in demography, economics, political science, public health, public policy, and sociology. Although the Social Security Administration has performed these forecasts since 1942, no systematic and comprehensive evaluation of their accuracy has ever been published.

Each year, the Office of the Chief Actuary of the Social Security Administration carries out the mandate for producing forecasts in the *Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance*

■ Konstantin Kashin is a PhD Candidate, Institute for Quantitative Social Science, Harvard University, Cambridge, Massachusetts. Gary King is Albert J. Weatherhead III University Professor, Institute for Quantitative Social Science, Harvard University, Cambridge, Massachusetts. Samir Soneji is Assistant Professor, Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, Lebanon, New Hampshire. Their email addresses are [kkashin@fas.harvard.edu](mailto:kkashin@fas.harvard.edu), [king@harvard.edu](mailto:king@harvard.edu), and [samir.soneji@dartmouth.edu](mailto:samir.soneji@dartmouth.edu).

<sup>†</sup>To access the Appendix and Data Appendix, visit <http://dx.doi.org/10.1257/jep.29.2.239>

*Trust Funds*, commonly known as the “Trustees Report.” Actuaries at the Social Security Administration (SSA) separately forecast demographic variables (for example, mortality rates) and economic variables (for example, labor force participation rates) that ultimately combine to produce solvency forecasts. In this article, we offer the first evaluation of Social Security forecasts that compares the SSA forecasts with observed truth; for example, we look at forecasts made in the 1980s, 1990s, and 2000s with outcomes that are now available. We do this first for demographic forecasts then for financial forecasts.<sup>1</sup>

Forecasts, of course, should not be expected to be precisely accurate. However, our analysis reveals several problems. First, Social Security Administration forecasting errors—as evaluated by how accurate the forecasts turned out to be—were approximately unbiased until 2000 and then became systematically biased afterward, and increasingly so over time. Second, most of the forecasting errors since 2000 are in the same direction, consistently misleading users of the forecasts to conclude that the Social Security Trust Funds are in better financial shape than turns out to be the case. Finally, the Social Security Administration’s informal uncertainty intervals appear to have become increasingly inaccurate since 2000. Although the Social Security Administration has recently begun to follow the recommendations of its panel of outside technical advisers on including certain types of more formal uncertainty estimates, a step that should be part of all government reporting (Manski 2013), these estimates have also not been systematically evaluated.

At present, the Office of the Chief Actuary, at the Social Security Administration, does not reveal in full how its forecasts are made and, as a result, no other person, party, or organization, in or out of government, has been able to make fully independent quantitative evaluations of policy proposals about Social Security. Even the Congressional Budget Office, which produces Social Security Trust Fund solvency forecasts, relies on the demographic forecasts produced by the Office of the Chief Actuary as inputs for its models. Thus, the Office of the Chief Actuary holds an unusual position within American politics of being the sole supplier of Social Security forecasts, as well as heading the only organization producing fully independent quantitative evaluations of policy proposals to alter Social Security. For each evaluation of a proposed policy, the Office of the Chief Actuary estimates the effect on key financial outcomes that assess the solvency of the Trust Funds. For the vast majority of policy proposals evaluated by the Office of the Chief Actuary, the estimated financial impact is smaller than almost all of SSA’s forecasting errors since 2000. Social Security Administration forecasts of current law and its counterfactual evaluation of policy proposals share the same growing bias because both are based on the same forecasting methodology. Additionally, the

<sup>1</sup> A replication dataset available with this article at <http://e-jep.org> and in Dataverse (Kashin, King, and Soneji 2015b) summarizes our data sources and all the information necessary to reproduce our results. An online Appendix, available at the same website, offers extra robustness checks.

uncertainty surrounding the estimated effects of proposed policies, which would likely be larger than the uncertainty in the forecasts under current law, usually dominate the estimated effect of the policy.

In the conclusion of the article, we argue that the Social Security Administration and its Office of the Chief Actuary should follow best practices in academia and many other parts of government and make their forecasting procedures public and replicable, and should calculate and report calibrated uncertainty intervals for all forecasts. In a companion paper, we offer an explanation for the origin of the biases reported here and propose simple structural ways of changing the system to fix the problems going forward (Kashin, King, and Soneji 2015a).

## **Demographic Forecasts**

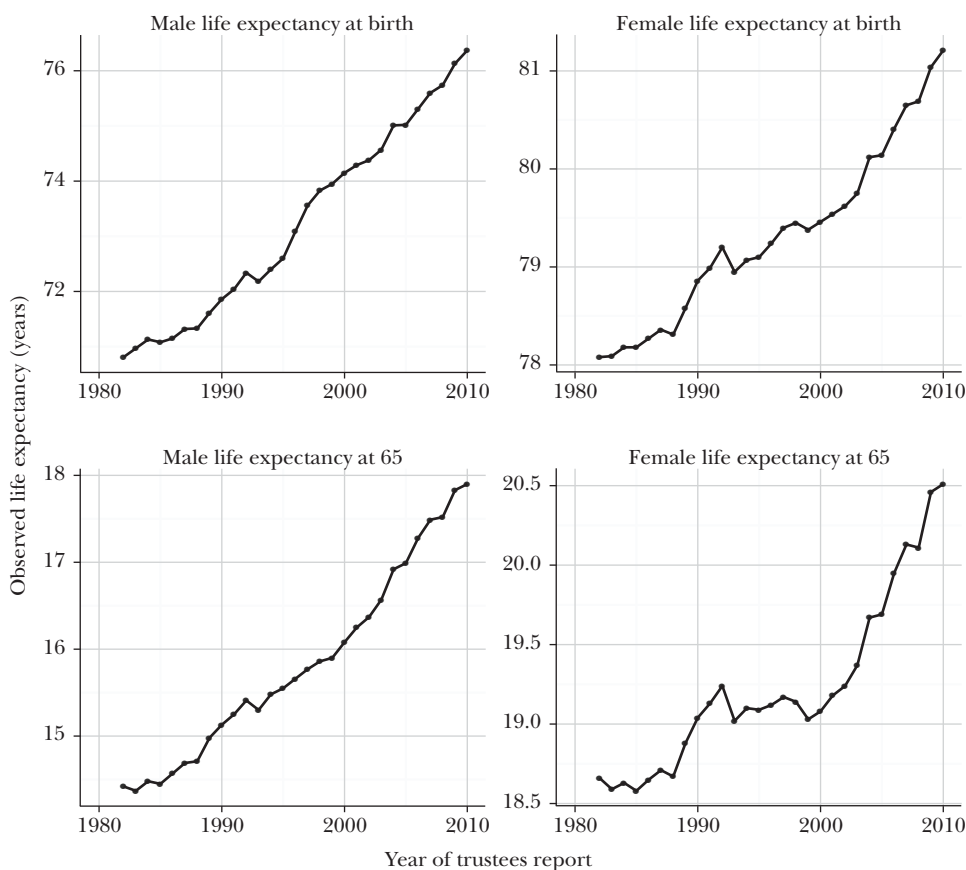
Demographic variables important to solvency forecasts from the Social Security Administration include mortality, fertility, and migration. Higher levels of fertility and migration increase the number of workers who contribute payroll taxes and increase long-term solvency. Lower levels of mortality, especially among those age 65 years and older, increase the number of retirees who receive benefits and decrease long-term solvency. Moreover, if Americans live longer than the forecasts predict, they will draw benefits for more years than expected and the Trust Funds will become exhausted sooner than anticipated. As Diamond and Orszag (2005, p. 63) explain, the increase in benefit payments from longer lives is not counterbalanced by an increase in payroll tax receipts because the system is designed to be approximately fair on average from an actuarial standpoint—the longer lives were not taken into account in what was paid into the system earlier in working life. People with longer working careers also receive higher benefits compared to those starting their careers at later ages.

### **Observed Demographic Data**

As a baseline, we present four observed time series in Figure 1. Life expectancy for males and females, both at birth and at age 65, are relatively smooth over almost the entire time period, as can be seen in all four graphs. Indeed, three of the four are approximately linear; the fourth, female life expectancy at 65, is not far from linear. The highly regular nature of these data suggests that relatively accurate forecasts should be possible.

There is a standard conceptual difficulty in measuring current life expectancy: How can the analyst describe life expectancy when people are still alive? We follow common practice here by using the concept of “period life expectancy.” This approach calculates life expectancy in a given year as the average number of years a person would expect to live if that person experienced the mortality rates in that given year over the course of a lifetime. Thus, life expectancy is a function of age-specific mortality rates and the average number of person-years contributed by those who die

*Figure 1*  
**Observed Period Life Expectancy**



*Note:* “Period life expectancy” for a year is a single-number summary of all the age-specific mortality rates for that same year and is interpreted as the average number of years a person could expect to live if he or she experienced the mortality rates of a given year over the course of their life.

at each age. The mortality rate for people of a given age equals the number of deaths in that age divided by the number of person-years lived in that age (the exposure).<sup>2</sup>

The Office of the Chief Actuary forecasts male and female life expectancy separately. The male and female population counts are then combined with sex-specific economic factors like estimates of the labor force participation rates, and sex-specific beneficiary rates like disability incidence rates to project the population of workers and beneficiaries. In turn, the number of male and female workers

<sup>2</sup> We use observed life expectancy based on Human Mortality Database life tables rather than the life tables from the Social Security Administration. The small differences in estimated life expectancy between the two sources do not account for the much larger error rates and patterns reported in this article. Both sources seek to estimate the conditional probability of death (and life expectancy, its single

and beneficiaries serve as inputs for predictions of the operations and actuarial status of the Trust Funds.

We begin in 1982, the earliest year with regular life expectancy forecasts from the Social Security Administration, and continue until 2010, the last year for which observed actual data have been released. For the years before 2001, the Social Security Administration only reveals information about its demographic forecast for years divisible by five. However, the observed time series are quite smooth and, hence, interpolations to other years should be accurate.

### **Forecasts**

We present our results in stages beginning in Figure 2 with an evaluation of forecasts for 2005 and 2010, the two years forecast by the largest number of Trustees Reports (with details for all years in our online Appendix). We compute forecast error (here and throughout) as the “intermediate scenario” forecast minus the observed value, so that positive values represent overestimates and negative values represent underestimates. The vertical axis is the forecast error for each of the four demographic variables, and the horizontal axis is the year of the Trustees Report when the forecast was made. Figure 2 illustrates four points.

First, despite the strong resemblance and very high correlation between male and female life expectancy in Figure 1, the forecast errors are substantially worse for males than females over most of the range of the forecasts. In some of the forecasts of the mid-1980s, the overestimate of female life expectancy is more-or-less offset by the underestimate of the male life expectancy, but in later years, both are underestimated.

Second, the patterns of error persist. For example, Figure 2 shows that every single Trustees Report for 23 years from 1982–2005 underestimated male life expectancy in 2005. Similarly, every forecast for 28 years from 1982–2010 underestimated male life expectancy in 2010.

Third, the forecasts for males do pass an obvious test by more closely approximating the truth as the year being forecast approaches. For females, errors have been smaller than for males until recently, but in the years from 2000–2005, when projecting female life expectancy at 65, forecast errors of female life expectancy actually increased as the year of the Trustees Report approached the year being forecast.

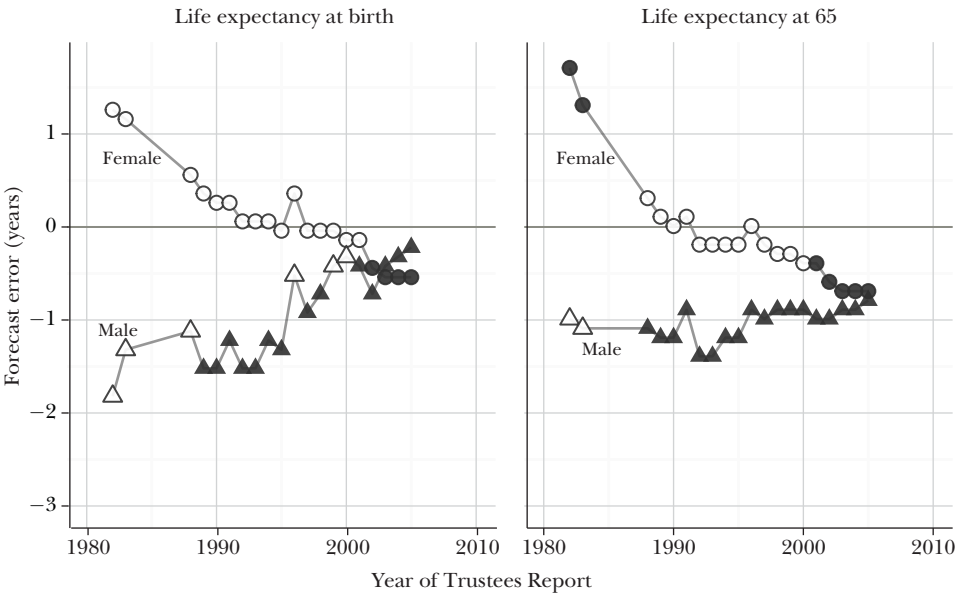
Fourth, a large number of the forecasts fall outside the uncertainty intervals offered by the Social Security Administration. In the forecasts, these uncertainty intervals are categorized as “high cost,” “intermediate cost,” and “low cost” scenarios. The high- and low-cost scenarios form the Social Security Administration uncertainty interval. In Figure 2, we color points white if the truth falls within

---

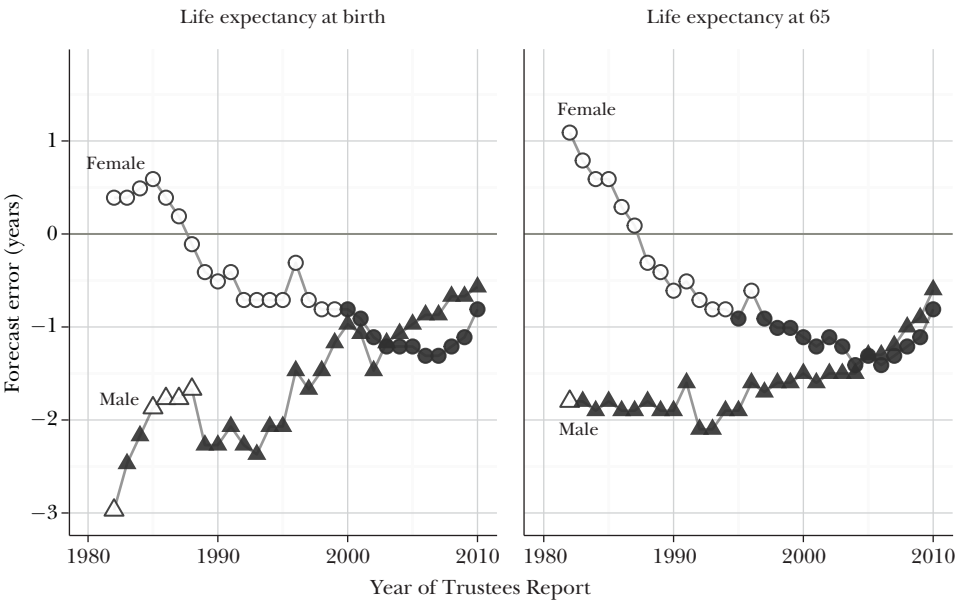
number summary), but the Human Mortality Database is the standard in the scientific literature for its emphasis on “comparability, flexibility, accessibility, [and] reproducibility,” for subjecting US Census counts and National Center for Health Statistics death counts to international quality standards, and for including all potential beneficiaries in the analyses, as discussed in an overview to the database (<http://www.mortality.org/Public/Overview.php>; accessed March 20, 2015). An online Appendix repeats all demographic analyses in this section with Social Security Administration data; no important differences arise compared with the results presented here.

*Figure 2*  
**Forecast Error of Life Expectancy in 2005 and 2010 by Year of Trustees Report**

A: 2005 Life Expectancy Forecast Error



B: 2010 Life Expectancy Forecast Error



*Note:* The circles (females) and triangles (males) are colored white when truth falls within Social Security Administration uncertainty intervals and colored black when the truth falls outside these uncertainty intervals.

these “uncertainty intervals” and black if the truth falls outside of these intervals. Although any forecast is of course uncertain and errors are to be expected, uncertainty intervals should still capture the truth with some known frequency. We find that only one of the 29 uncertainty intervals for male life expectancy at age 65 for 2010 actually captured the true outcome. In the years after 2000, every single forecast for year 2010 male and female life expectancy at birth and at age 65 was underestimated, and the true outcome fell outside the uncertainty intervals.

The uncertainty intervals reported by the Social Security Administration are given no formal statistical basis in published materials, and thus we tried to assess how these intervals were qualitatively presented. In Trustees Reports from the earlier part of our period, the early and mid-1980s, the Social Security Administration wrote about the intervals as one would discuss wide confidence intervals, perhaps at a 90 percent confidence interval, and readers were warned that the confidence intervals might not necessarily cover the truth. In recent years, especially after 2000, the Trustees Reports became more confident in these intervals. Since 2003, the Trustees Reports have included an appendix referring to a stochastic model that attempts to formalize the uncertainty of their forecasts. The model itself is not publicly available, so outside analysts cannot evaluate how it has been calibrated or evaluated. But the Trustees describe the uncertainty intervals in qualitative terms that one would typically use to discuss something stronger than a 95 percent confidence level. For example, in the 2011, 2012, and 2013 Trustees Reports, the report repeated the same definition: “In the future, the costs of OASI, DI, and the combined OASDI programs as a percentage of taxable payroll are unlikely to fall outside the range encompassed by alternatives I [low cost] and III [high cost] because alternatives I and III define a wide range of demographic and economic conditions.” (OASI refers to the old age and survivors program, which is commonly known as Social Security, while DI refers to the disability insurance program.)

In short, the post-2000 forecasts all indicated that both men and women would have lived shorter lives than they did and also offered uncertainty ranges implying that the Trust Funds were on firmer financial ground than turned out to be warranted. We reach an identical conclusion when we examine the forecast error over all Trustees Reports for all observed years, as shown in the online Appendix.

### **Uncertainty Intervals**

Finally, we analyze the set of all the Social Security Administration life expectancy forecasts with respect to uncertainty interval coverage. Figure 3 plots the year of the Trustees Report (horizontally) by the year of the forecast (vertically), with one square for each forecast colored white when the truth fell within the uncertainty interval and black when the truth fell outside the interval.

The results in Figure 3 demonstrate systematic problems with the uncertainty intervals used by the Social Security Administration. The uncertainty intervals failed to capture the truth for almost every forecast made since 2000 for all four demographic variables. For the graphs on male life expectancy at birth and at age 65 (the two graphs on the left), the problem began approximately in 1990.



Figure 3

**Uncertainty Interval Coverage by Year of Trustees Report and Year of Forecast**

Notes: White indicates uncertainty interval covered the truth, black indicates that it did not, and gray “X” indicates that the Social Security Administration did not provide an uncertainty interval. Contemporaneous forecast error is possible because of the time lag (typically three to four years) in finalizing mortality data.

We might expect that some uncertainty intervals fail to capture the eventually observed truth, especially when the forecast was made many years earlier than the year forecast. But since about 2000, the uncertainty intervals consistently failed to capture the truth for male and female life expectancy at birth and age 65. Apparently, the Social Security Administration did not perform a correction if and when these systematic errors became known.

**Forecast Biases Are Not Explained by the Great Recession**

Could the systematic forecasting biases documented in this section be caused to some extent by the Great Recession, which lasted from December 2007 to June 2009? Historically, increases in unemployment have led to lower mortality primarily because of fewer accidental deaths (like road traffic fatalities) (Granados 2005; Stuckler, Meissner, Fishback, Basu, and McKee 2011) not counterbalanced by a small increase in the comparatively fewer number of suicides. Thus, a lengthy recession could potentially explain life expectancies longer than predicted. But although the recession may explain some of the forecasting error, it cannot explain most of it.

First, the Great Recession began in December 2007, when the Social Security Administration had already been underestimating Americans’ life expectancy for several years prior. Second, the mortality data and errors in forecasting mortality

from one year to the next are relatively smooth functions of time—that is, the errors do not increase when the recession arrived. Finally, the Great Recession cannot account for the 0.6-year forecast error in male life expectancy and 0.8-year forecast error in female life expectancy made by the Social Security Administration in 2010. During the 18-month recession, unemployment increased 4.6 percentage points from a trough of 4.9 percent in February 2008 to a peak of 9.5 percent in June 2009. Previous US- and European-based studies estimate that mortality rates decline approximately 0.5 percent for every 1 percent in unemployment (Ruhm 2000). Thus, the 4.6 percentage point increase in employment during the Great Recession would correspond approximately to a 2.3 percent decline in mortality rates. For comparison, the inaccuracies in projected male and female life expectancies correspond to a 5.2 and 7.6 percent decline in mortality rates, respectively.

### **A Note about Fertility and Immigration**

We also evaluated the performance of Social Security Administration forecasts of fertility and migration (with results shown in the online Appendix). Recent forecasts of the total fertility rate exhibited persistent and growing error, and the forecasts were overly confident. For example, the error in forecasts of the total fertility rate in 2010 grew—rather than shrank—across successive Trustees Reports. The forecast error of the 2010 total fertility rate in the 2010 Trustees Report was 0.15, which translated to approximately 315,000 more births forecasted than actually occurred (8 percent of total births in 2010).<sup>3</sup> As with mortality, forecast biases in fertility are not explained by the Great Recession. The rise in unemployment during the Great Recession led to a fertility decline of approximately 5 percent. Yet, the inaccuracy in the total fertility rate forecasted in 2010 corresponded to an approximately 8 percent difference in fertility rates. Overall, the forecast error in fertility makes the US population seemingly younger than it really is and, consequently, the Social Security Trust Funds healthier than they may be.

In contrast to mortality and fertility, Social Security Administration forecasts of legal immigration (the largest component of overall immigration) were far less biased and the confidence intervals seem appropriate. For example, the error in the forecast of net legal immigration in 2010 declined across successive Trustees Reports. By the 2010 Trustees Report, the forecast error was less than 1 percent of the observed number of net legal immigrants in 2010.

The results of mortality, fertility, and immigration forecasts may illuminate some of the reasons why the Social Security Administration varies in its performance of forecasting these three demographic components. As we discuss in detail in Kashin, King, and Soneji (2015a), a constellation of factors may have interacted to produce biased mortality and fertility forecasts. First, the forecasting method itself allows for the introduction of unintentional bias because it apparently involves a very large number (previously 210, now 150) of interrelated subjective decisions about rates

<sup>3</sup> The 2010 Trustees Report included historical fertility up to 2006 because of the time lag in reporting final birth data.

of mortality decline. Second, as Social Security reform has become more politically charged, the Social Security Administration seems to have disregarded the continued advice of its outside technical advisers to assume a more rapid increase in life expectancy. Third, mortality rates decreased at an ever faster pace after about 2000, but the Social Security Administration mortality forecasts did not keep pace with this change in input data.

Some of the same factors that possibly produce biased mortality forecasts may occur for fertility, too. The Social Security Administration forecasting method for fertility also involves subjective decisions about future levels of fertility rates. In contrast to mortality and fertility, the level of legal immigration is annually set by Congress. The Social Security Administration forecast of net legal immigration largely follows this Congressionally-set level; at present, it sets the ratio of legal emigration to legal immigration to 25 percent.

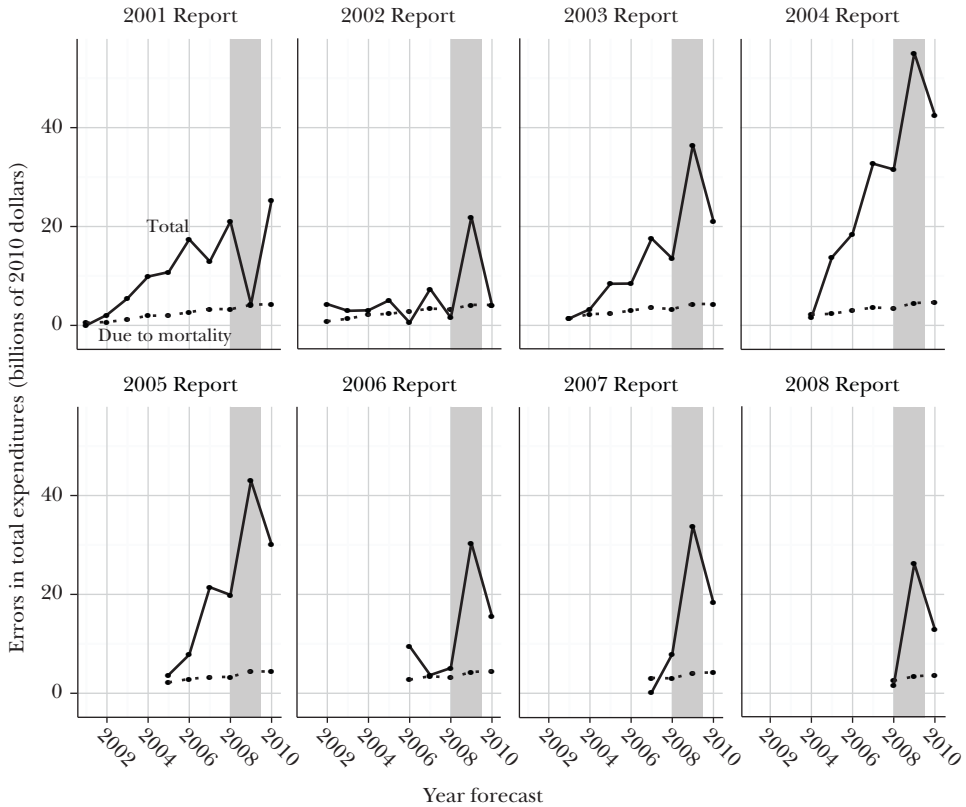
## Financial Forecasts

We next consider Social Security Administration forecasts of Trust Fund solvency, for which demographic forecasts serve as a key input. In particular, we examine forecasts and observed outcomes of the three most commonly cited financial indicators when discussing the health of Social Security: the *cost rate*, the *trust fund balance*, and the *trust fund ratio*. The cost rate equals the overall cost of the Social Security program in a given year divided by the taxable payroll for that year. The trust fund balance equals the difference between projected annual income and projected annual cost, as a percentage of the taxable payroll. The trust fund ratio equals the assets of the Social Security Trust Funds at the beginning of a calendar year divided by the expected expenditure for that year.

We collect all forecasts for each measure published in the annual Trustees Reports from 1978, when the reports began consistent reporting of financial indicators, until 2013. The reports usually include yearly forecasts between the year of the report and 10 years in the future and then every fifth subsequent year. After 2000, single-year supplemental tables are available online.

These three financial indicators directly relate to the economic and public policy debates that have occurred over nearly the entire lifetime of Social Security. After the Social Security Amendments of 1983, for example, the trust fund balance increased primarily because of higher payroll tax rates intended to build up a surplus in the trust fund, although benefit levels increased, too. Numerous economic studies find Social Security affects personal savings through reduction of disposable income because of payroll taxes and anticipated benefits during retirement (Harris 1941; Feldstein 1974; Diamond and Hausman 1984). Gramlich (1996) argued that proposed Social Security reform faces competing challenges in political economy: ensuring long-run actuarial balance while not lowering the ratio of discounted benefits to discounted taxes paid (the “money’s worth” ratio). The long-run actuarial balance, a function of the trust fund balance and cost rate, can

Figure 4

**Cost of Mortality Forecasting Errors**

*Notes:* Each panel of the figure corresponds to a Trustees Report. Within each panel, we plot the forecast error in total Social Security expenditures (solid lines) and the forecast error in total Social Security expenditures due to mortality forecasting errors (dashed lines). Finally, we represent the Great Recession as a vertical grey region.

be maintained by raising payroll tax rates or lowering benefit levels, although such changes would reduce the money's worth ratio. Other proposed reforms, such as individual accounts and personal savings accounts, offer a possibility of extending the solvency of Social Security and maintaining the money's worth ratio but face intense public scrutiny (Samwick 1999).

**The Cost of Mortality Forecasting Errors**

Before turning to the three financial indicators, we begin by comparing the forecast errors in cost that are specifically due to forecast errors in mortality (dashed line, Figure 4) to the overall forecast errors in cost (solid line). In theory, either of these forecast errors in cost could be larger than the other because forecast errors in cost potentially come from many inputs other than mortality.

For each Trustees Report and forecast year, we estimate the number of additional retirees that the Social Security Administration did not expect because of errors in predicting life expectancy. For example, the 2005 Trustees Report under-forecasted male life expectancy at age 65 in the year 2010 by 1.3 years (forecast: 16.6 years; true outcome: 17.9 years). The 1.3 years under-forecast of life expectancy corresponds to approximately 151,000 male beneficiaries. We estimate the forecast errors in costs due to forecast error in mortality as the product of the total number of additional beneficiaries and the average benefit amount per year. For this figure, we plot the forecast year on the horizontal axis and the forecast error in cost (in billions of 2010 dollars) on the vertical axis. Each panel presents forecasts from different Trustees Reports. The years of the Great Recession are denoted by the grey shaded region. To put the figure into perspective, the total cost of the Social Security program in 2010 was \$712.5 billion.

Figure 4 emphasizes four points. First, mortality is a highly predictable part of the overall forecast error in cost, as evidenced by the highly smooth and almost linear dashed lines in each panel. Second, for many years, forecast errors in cost specifically due to forecast errors in mortality were a large fraction of the overall forecast error in cost. Third, the forecast errors in cost due to forecast errors in mortality are neither random nor constant. The errors increase secularly and thus strongly suggest the existence of information that can be used to improve forecasting performance. Finally, the overall forecast errors in costs are highly variable relative to errors due to mortality. They are much larger during the Great Recession, shown by the vertical shaded area, but these overall forecast errors in costs were also large at times well before the onset of the Great Recession.

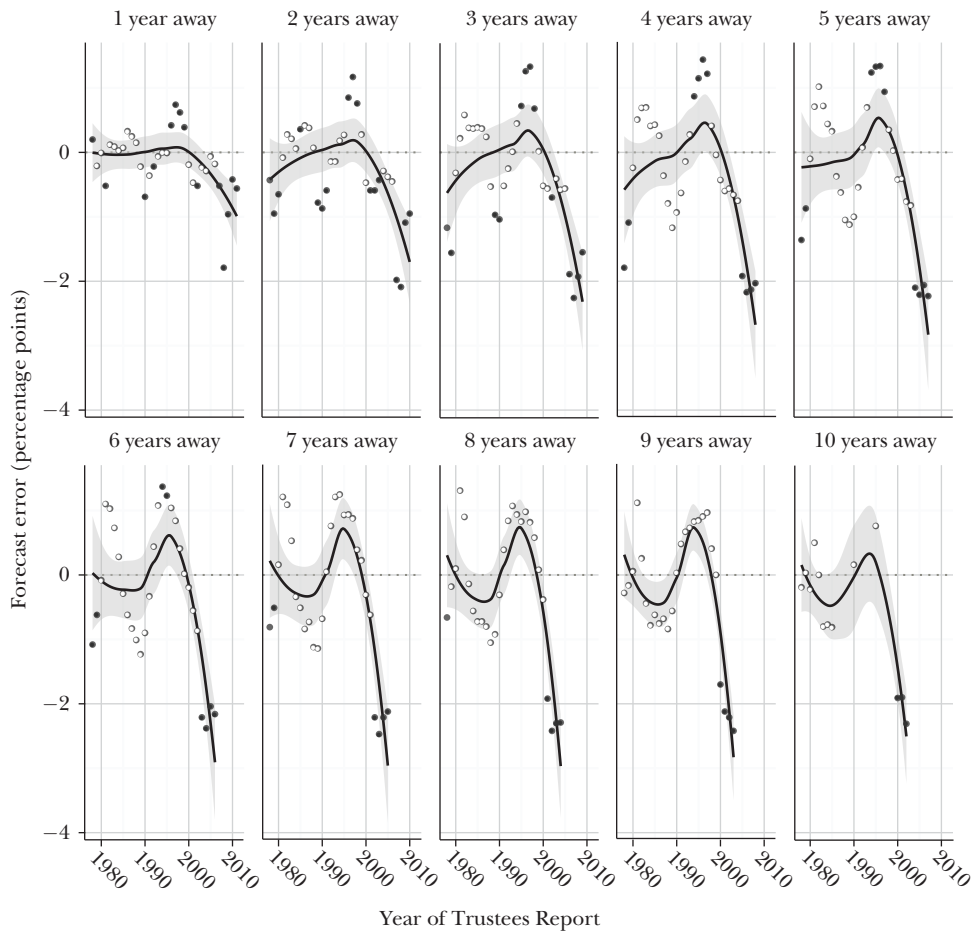
### Cost Rate Forecasting Errors

Figure 5 reports the forecast error in the cost rate (the vertical axis in each panel) made in a Trustees Report in the given year (the horizontal axis in each panel) for a number of years out into the future (in the title of each panel). For example, the upper-left panel shows the forecast error in the cost rate for forecasts made one year in advance of the year forecast. A value of zero on the vertical axis indicates that the forecast was perfectly accurate. White points fall within the forecast uncertainty interval and black points fall outside. To enhance readability, we superimpose on each panel a smoothed line showing the path of the errors.<sup>4</sup>

The pattern of forecast errors in Figure 5 is striking. Forecasts from Trustees Reports until about 2000 were approximately unbiased, which can be seen by the roughly random scatter of points vertically around the horizontal line at zero forecast error. However, forecasts from Trustees Reports after roughly the year 2000 were increasingly biased over time, and all in the same direction. Congress and

<sup>4</sup> The smoothed line is estimated with a locally weighted scatterplot smoothing (LOESS) procedure, in which the predicted error  $t$  for Trustees Report is calculated based on a local polynomial of degree 2, fit to neighboring observations. These observations are weighted by their tricubic distance from the Trustees Report in question.

Figure 5

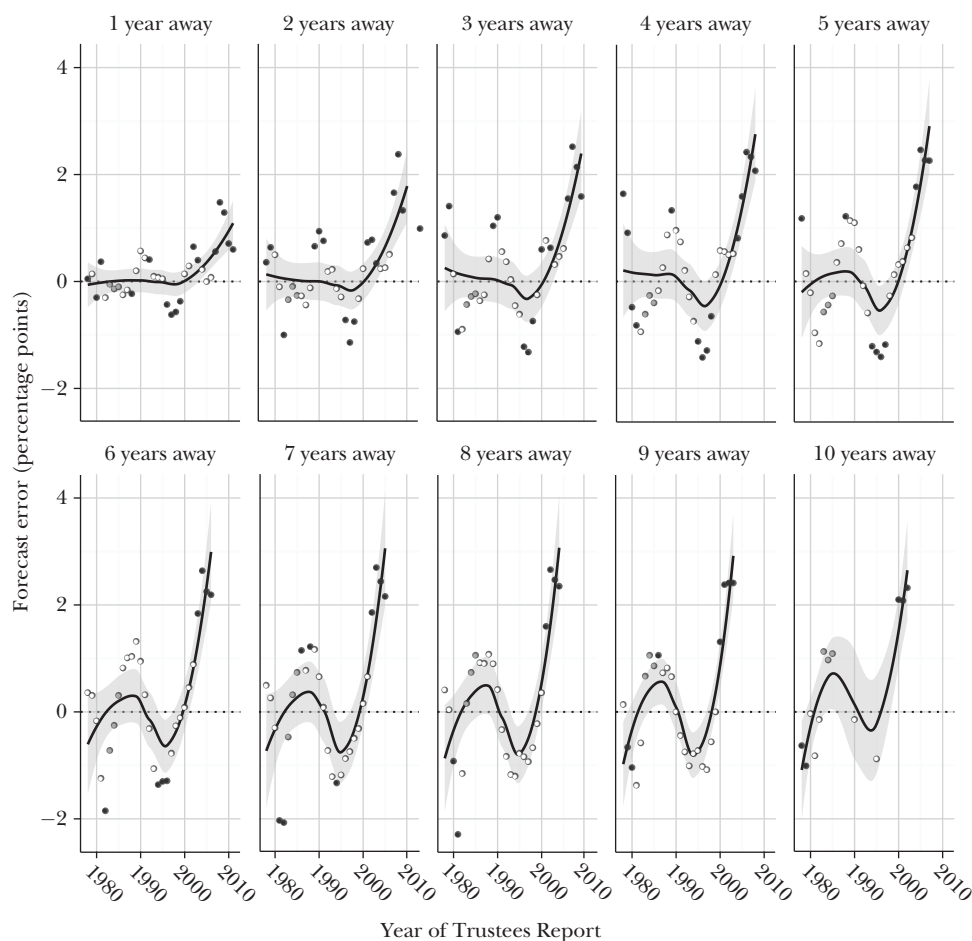
**Cost Rate Forecasting Errors**

*Notes:* The graphs show forecast errors in the cost rate (vertically) by the year of the forecast (horizontally) by how many years into the future the forecast is made (in the title of each panel). Cost rate forecasting errors are overestimates if positive and underestimates if negative. Points are white if the error is within the Social Security Administration's uncertainty interval and black otherwise. To enhance readability, we superimpose on each panel a smoothed line showing the path of the errors. This error bar around this line relates to the observed path of the errors and not to the SSA predicted path of observations. See footnote 4.

other users of these forecasts would have been misled into thinking that the cost of the Social Security program was less than it actually turned out to be. This is as true for forecasts one year into the future (top left) as for forecasts 10 years into the future (bottom right). As expected, the errors are larger for forecasts farther into the future.

Finally, Figure 5 shows that the largest errors are also most likely to be outside the uncertainty intervals (as indicated by black dots). The purpose of uncertainty

Figure 6

**Trust Fund Balance Forecasting Errors**

*Notes:* The graphs show forecast errors in balance (vertically) by the year of the forecast (horizontally) by how many years into the future the forecast is made (in the title of each panel). Positive errors overestimate the Trust Fund balance; negative errors underestimate it. Points are white if the error is within the Social Security Administration's uncertainty interval and black otherwise. To enhance readability, we superimpose on each panel a smoothed line showing the path of the errors. This error bar around this line relates to the path of the observed errors and not to the SSA predicted path of observations. See footnote 4.

estimates is to protect oneself from drawing overconfident conclusions from the data, and if estimates are consistently falling outside those uncertainty intervals, then alteration and improvement in the forecasting process should follow.

**Trust Fund Balance Forecasting Errors**

A positive annual trust fund balance indicates the program has a surplus for the year and a negative trust fund balance translates to a deficit. We present Figure 6 in

the same format as Figure 5. The evaluation of forecasting errors in the trust fund balance leads us to the same conclusions as forecasting errors in the cost rate. The Social Security Administration forecasts of trust fund balances were approximately unbiased until about 2000, after which they become substantially biased. Moreover, the direction of the biases are all in the same direction, making the Social Security trust funds look healthier than they turned out to be. The reported uncertainty intervals are again overconfident.

### **Trust Fund Ratio Forecasting Errors**

When the trust fund ratio equals 0 percent or becomes negative, the Social Security Trust Funds are insolvent. The Trust Funds are deemed financially adequate in the short term if the ratio stays above 100 percent for the first 10 forecasted years. Insolvency does not release the federal government from its obligation to pay some level of benefits to qualified individuals (Meyerson 2014). The Social Security Act stipulates that every fully insured individual is entitled to receive benefits. On the other side, the Antideficiency Act prohibits the federal government from paying Social Security benefits beyond the balance of the Trust Funds. Once insolvency occurs, beneficiaries would either receive delayed or lower benefit payments.

In Figure 7, we present results in a form parallel to Figures 5 and 6. While the uncertainty intervals appear to have better coverage when compared to the cost rate and trust fund balance metrics, the overall results in this figure confirm the main results from our analysis of the cost rate and trust fund balance. First, trust fund ratio forecast errors are approximately unbiased from 1978 through about the year 2000, as indicated by the dots scattered randomly above and below the vertical line drawn at zero. After 2000, forecast errors became increasingly biased, and in the same direction. Trustees Reports after 2000 all overestimated the assets in the program and overestimated solvency of the Trust Funds. The size of this bias has increased over time, with the more recent Trustee Reports being less and less reliable. Finally, the coverage of uncertainty estimates did not improve over time and were strongly and positively correlated with the size of the absolute error.

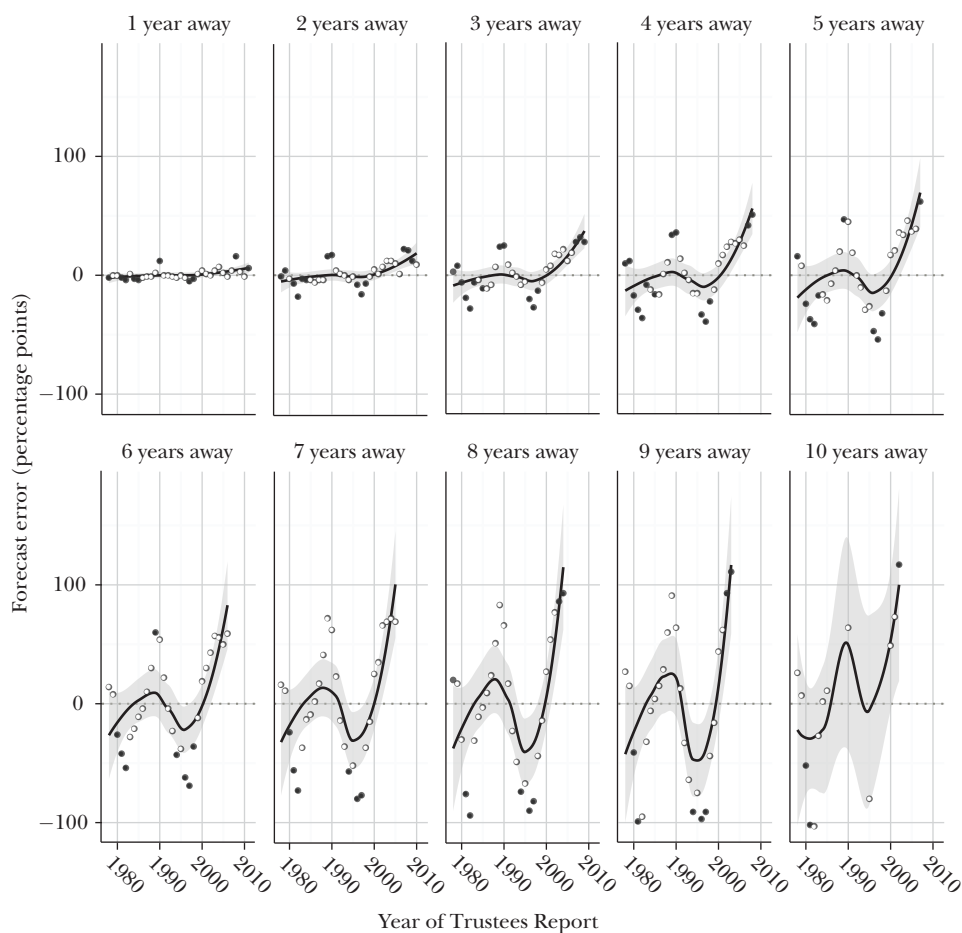
### **Implication of Financial Forecasting Errors for Proposal Scoring**

In addition to producing the annual Trustees Report, the Office of the Chief Actuary also scores policy proposals to alter Social Security submitted by members of Congress, the administration, and select professional organizations. For each of the policy proposals it scores, the Office of the Chief Actuary makes point-estimate predictions about what would happen to one or more financial metrics, such as those we study above, if the proposal became law. Although the Office of the Chief Actuary includes no uncertainty measures with these predictions, we can estimate their uncertainty on the basis of our evaluation of their forecasts.

The extent of uncertainty in these counterfactual predictions can be divided into two components. The first is the inherent uncertainty of the effect of the intervention if the law changes as proposed. The second is the uncertainty in forecasting



Figure 7

**Trust Fund Ratio Forecasting Errors**

*Notes:* The graphs show forecast errors in the trust fund ratio (vertically) by the year of the Trustees Report forecast (horizontally) by how many years into the future the forecast is made (in the title of each panel). Positive errors overestimate the Trust Fund ratio; negative errors underestimate it. Points are white if the error is within the Social Security Administration's uncertainty interval and black otherwise. To enhance readability, we superimpose on each panel a smoothed line showing the path of the errors. This error bar around this line relates to the path of the observed errors and not to the SSA predicted path of observations. See footnote 4.

the same financial indicators under current law, as we do earlier. We use our evaluation of the second component as a lower bound for the uncertainty of the Office of the Chief Actuary's policy scoring.

The Social Security Administration evaluated 93 proposals since 2000, which resulted in 110 assessments of financial indicators for which we can evaluate

forecasting performance (see <http://ssa.gov/oact/solvency>). For example, in 2015, the Social Security Administration evaluated the effect of President Obama's Executive Actions for immigration on Social Security solvency. The Chief Actuary concluded immigration reform would increase the cost rate by an average 0.04 percent over the next 75 years, which is considerably smaller than most cost rate forecasting errors made since 2000. Overall, we found that 42 percent of policy assessments by the Office of the Chief Actuary predicted changes in Social Security finances that were smaller than the average forecasting error made since 2000. And 95 percent of the assessments concluded by predicting changes in Social Security finances that were smaller than the maximum forecasting error made since 2000.

Members of Congress and the public devote considerable energy debating policy proposals on the basis of these evaluations. Presidents and their opponents tout the merits of policy proposals to engender public support. But if the lower bound on the magnitude of forecasting errors exceeds the estimated effect of the reforms, then it seems likely these discussions and debates are not grounded in the best information available.

## **Conclusions and Recommendations**

In recent years, especially after about 2000, the Social Security Administration began issuing systematically biased forecasts with overconfident assessments of uncertainty. Reliance on such forecasts led policymakers and other users of the forecasts to conclude that the Social Security Trust Funds were on firmer financial ground than actually turned out to be the case. We focus here on three steps that the Social Security Administration should take to ensure this problem is addressed; other suggestions are offered in King, Kashin, and Soneji (2015a).

First, forecasting mistakes are no embarrassment unless the forecaster fails to learn from them. Thus, we recommend that the Social Security Administration publish annually a systematic and comprehensive evaluation of its forecasting performance for both demographic factors and financial solvency. This best practice of forecasting self-evaluation is routine among academic researchers (Lee and Miller 2001) and professional actuaries (Lu and Won 2011), for social security programs in other countries (Shaw 2007), and in other parts of the US government, such as the Congressional Budget Office (2013), the Census Bureau (Wang 2002), the Bureau of Labor Statistics (Wyatt 2010), and even other parts of the Social Security Administration itself (US Social Security Administration's "Fiscal Year 2013 Major Evaluations" and *Social Security Administration Agency Financial Report: Fiscal Year 2013*). Every future Trustees Report, without exception, should include a routine evaluation of all prior forecasts, and a discussion of what forecasting mistakes were made, what was learned from the mistakes, and what actions might be taken to improve forecasts going forward.

Second, the Social Security Administration withholds from public view much of the data and procedures it uses to make many of its forecasts. The Office of the

Chief Actuary, which produces the demographic and economic forecasts, does not share much of its data and procedures even with other parts of the Social Security Administration. Currently, the best anyone can do to understand how the Social Security Administration forecasts work is to attempt to reverse-engineer their results (as done by many involved in the policy process and authors of simulation programs such as SSASIM by the Policy Simulation Group; see also King and Soneji 2011, and Soneji and King 2012). The “replication standard” for data sharing is the widely understood and accepted best practice throughout the scientific community (King 1995, p. 444) and echoed in the Obama administration’s executive orders requiring “a presumption in favor of openness,” and that data produced by government be “accessible, discoverable, and usable by the public” (Executive Office of the President, Memorandum, May 9, 2013).

Finally, it appears to us and to other outside observers that the forecasting procedures used by the Social Security Administration are in a number of ways ad hoc and suboptimal. These approaches also fail to take advantage of many of the dramatic improvements in statistical modeling over the last several decades (for example, Girosi and King 2008; King and Soneji 2011). Even some explicitly quantitative parts of the methods seem idiosyncratic or unnecessarily model dependent.

Our study reveals systematic errors in both demographic and solvency forecasts. Forecasting errors in economic variables, such as the labor force participation rate and growth in average wages, may also contribute to systematic errors in Trust Fund solvency forecasts. For the disability program of Social Security, forecasting errors in the disability incidence rate may be an additional important source of solvency forecast error.

This list of “best practices” is neither new nor controversial. There is a Social Security Advisory Board Technical Panels on Assumptions and Methods made up of outside experts. The Social Security Administration’s own outside advisors have repeatedly and emphatically recommended that the Office of the Chief Actuary make its data and procedures widely available, and allow its work to be replicated by outside groups. The collective efforts of the scientific community could easily be marshaled to improve the difficult forecasting task that confronts the Social Security Administration, all essentially without cost to the taxpayer. The creation of transparent forecasting procedures will also enable members of Congress and partisans on all sides to consider alternative assumptions explicitly when they debate proposals to ensure the solvency of Social Security. Forecasts of Social Security solvency also shape debates on immigration, public health, taxation, and income redistribution from working age adults to retirees. Accurate forecasts would help ensure these debates are based on the best information available.

■ *For helpful advice or comments, we are grateful to Bill Alpert, Jim Alt, Steve Ansolabehere, Neal Beck, Nicholas Christakis, Mo Fiorina, Dan Gilbert, Alexander Hertel-Fernandez, Martin Holmer, David Langer, and Theda Skocpol. We received approval from our university Institutional Review Boards for this study.*

## References

- Congressional Budget Office, US Congress.** 2013. "CBO's Economic Forecasting Record: 2013 Update." January.
- Diamond, Peter, and Jerry Hausman.** 1984. "Individual Retirement and Savings Behavior." *Journal of Public Economics* 23(1–2): 81–114.
- Diamond, Peter A., and Peter R. Orszag.** 2005. *Saving Social Security: A Balanced Approach*. Revised edition (paperback). Brookings Institution Press.
- Executive Office of the President, Office of Management and Budget.** 2013. "Memorandum for the Heads of Executive Departments and Agencies." M-13-13. May 9. <https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>.
- Feldstein, Martin.** 1974. "Social Security, Induced Retirement, and Aggregate Capital Accumulation." *Journal of Political Economy* 82(5): 905–26.
- Girosi, Federico, and Gary King.** 2008. *Demographic Forecasting*. Princeton University Press.
- Gramlich, Edward M.** 1996. "Different Approaches for Dealing with Social Security." *Journal of Economic Perspectives* 10(3): 55–66.
- Granados, José A. Tapia.** 2005. "Increasing Mortality during the Expansions of the US Economy, 1900–1996." *International Journal of Epidemiology* 34(6): 1194–1202.
- Harris, Seymour E.** 1941. *Economics of Social Security*. New York: McGraw-Hill.
- Human Mortality Database.** N.d. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org) (data downloaded on June 4, 2014).
- Kashin, Konstantin, Gary King, and Samir Soneji.** 2015a. "Explaining Nontransparency and Increasing Bias in U.S. Social Security Administration Forecasts." <http://j.mp/SSAbias>. (Forthcoming in *Political Analysis*.)
- Kashin, Konstantin, Gary King, and Samir Soneji.** 2015b. "Replication Data for: Systematic Bias and Nontransparency in US Social Security Administration Forecasts." UNF:5:1oerGFXQ0Bu9bcMFU5/t2A== <http://dx.doi.org/10.7910/DVN/28122>. Harvard Dataverse [Distributor] V1 [Version].
- King, Gary.** 1995. "Replication, Replication." *PS: Political Science and Politics* 28(3): 443–99. <http://j.mp/jCyf1>.
- King, Gary, and Samir Soneji.** 2011. "The Future of Death in America." *Demographic Research* 25(1): 1–38. <http://j.mp/iXUpBv>.
- Lee, Ronald, and Timothy Miller.** 2001. "Evaluating the Performance of the Lee–Carter Method for Forecasting Mortality." *Demography* 38(4): 537–49.
- Lu, Joseph, and Wun Wong.** 2011. "Mortality Improvement in the USA: Analysis, Projections and Extreme Scenarios." Technical report, Society of Actuaries Schaumburg, IL.
- Manski, Charles F.** 2013. *Public Policy in an Uncertain World: Analysis and Decisions*. Harvard University Press.
- Meyerson, Noah P.** 2014. *Social Security: What Would Happen If the Trust Funds Ran Out?* Technical report. Federal Publications, Key Workplace Documents at DigitalCommons@ILR.
- Ruhm, Christopher J.** 2000. "Are Recessions Good for Your Health?" *Quarterly Journal of Economics* 115(2): 617–650.
- Samwick, Andrew A.** 1999. "Social Security Reform in the United States." *National Tax Journal* 52(4): 819–842.
- Shaw, Chris.** 2007. "Fifty Years of United Kingdom National Population Projections: How Accurate Have They Been?" *Population Trends* 128: 8–23.
- Soneji, Samir, and Gary King.** 2012. "Statistical Security for Social Security." *Demography* 49(3): 1037–1060. <http://j.mp/Qv1a7N>.
- Stuckler, David, Christopher Meissner, Price Fishback, Sanjay Basu, and Martin McKee.** 2011. "Banking Crises and Mortality during the Great Depression: Evidence from US Urban Populations, 1929–1937." *Journal of Epidemiology and Community Health* 66(5): 410–19.
- Wang, Ching-li.** 2002. "Evaluation of Census Bureau's 1995–2025 State Population Projections." Working Paper No. POP-WP067, US Census Bureau.
- Wyatt, Ian D.** 2010. "Evaluating the 1996–2006 Employment Projections." *Monthly Labor Review*, September.
- US Social Security Administration.** 2012. "Fiscal Year 2013 Major Evaluations." Charts. <http://j.mp/SSAevals>.
- US Social Security Administration.** 2013. *Social Security Administration Agency Financial Report: Fiscal Year 2013*.
- US Social Security Administration.** Various Years. *Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds*.



**This article has been cited by:**

1. James Gorry, Dean Scrimgeour. 2018. USING ENGEL CURVES TO ESTIMATE CONSUMER PRICE INDEX BIAS FOR THE ELDERLY. *Contemporary Economic Policy* **36**:3, 539-553. [[Crossref](#)]
2. Carlos Patrick Alves da Silva, Claudio Alberto Castelo Branco Puty, Marcelino Silva da Silva, Solon Venâncio de Carvalho, Carlos Renato Lisboa Francês. 2017. Financial forecasts accuracy in Brazil's social security system. *PLOS ONE* **12**:8, e0184353. [[Crossref](#)]
3. Nicholas Diakopoulos, Oliver Deussen. 2017. Brauchen wir eine Rechenschaftspflicht für algorithmische Entscheidungen?. *Informatik-Spektrum* **40**:4, 362-366. [[Crossref](#)]
4. Marijn Janssen, George Kuk. 2016. The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly* **33**:3, 371-377. [[Crossref](#)]
5. David D. Jones. Issues in Applying Discount Rates 135-144. [[Crossref](#)]
6. A. Alesina, A. Passalacqua. The Political Economy of Government Debt 2599-2651. [[Crossref](#)]