

Systematic Bias and Nontransparency in US Social Security Administration Forecasts: Online Appendix

Konstantin Kashin*

Gary King†

Samir Soneji‡

March 2015

Contents

1	Data Sources	2
2	Main Figures in Color	4
3	Demographic Results Using SSA Mortality Data	11
4	Life Expectancy Forecast Errors for 1985–2010	14
5	Total Fertility Rate Forecast Errors	16
6	Migration Forecast Errors	18

*Ph.D. Candidate, Institute for Quantitative Social Science, Harvard University, Cambridge MA 02138; konstantinkashin.com, kvkashin@gmail.com, (617) 871-9884.

†Albert J. Weatherhead III University Professor, Institute for Quantitative Social Science, 1737 Cambridge Street, Harvard University, Cambridge MA 02138; GaryKing.org, king@harvard.edu, (617) 500-7570.

‡Assistant Professor, The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, 35 Centerra Parkway, Lebanon, NH 03766; samir.soneji@dartmouth.edu

This is the Online Appendix to [Kashin, King and Soneji \(2015b\)](#). Replication data for the paper and the Appendix is available in [Kashin, King and Soneji \(2015a\)](#).

1 Data Sources

Demographics We obtain observed period life expectancy data for 1982–2010 from the Human Mortality Database (HMD, [mortality.org](#)). We collected all life expectancy forecasts published in the annual Trustees Reports 1982–2010. In reports prior to 2001, SSA published life expectancy at birth and at age 65 forecasts for males and females projected in 5-year intervals for a total of 75 years into the future. Post-2001, supplementary single-year tables are included online. Our sources are Tables 11 of Trustees Reports 1982-1991, Table II.D.2 of Trustees Reports 1992-2000, and Table V.A3 of the supplemental single-year tables of Trustees Reports 2001-2010. We calculate residuals as the difference between SSA’s “best guess” projection (intermediate-cost scenario / alternative II) and the observed life expectancy from HMD. To calculate the uncertainty interval, we use the minimum and maximum values of projected life expectancy across the three scenarios.

We obtain observed total fertility rate (TFR) data for 1982–2010 from the Human Fertility Database (HFD, [humanfertility.org](#)). We collected all TFR forecasts published in the annual Trustees Reports 1982–2010. In reports prior to 2001, SSA published TFR forecasts in 5-year intervals for a total of 75 years into the future in the same tables as life expectancy (see preceding paragraph). For 2001 and onwards, supplementary single-year Table V.A1 of each Trustees Report contains TFR forecasts. We calculate residuals and uncertainty intervals in an analogous manner to life expectancy.

Finally, we collected all forecasts of net legal immigration for 2005 and 2010 published in the 2000–2010 Trustees Reports, available on page 62 of the 2000 Trustees Report and in Table V.A1 of the 2001–2010 Trustees Reports. The observed levels of net legal immigration are available in Table V.A1 of the 2014 Trustees Report.

Financials For 1978–2012, we take the observed cost rate and balance from Table IV.B1 and the observed trust fund ratio from Table IV.B3 of the 2013 Trustees Report ([j.mp/2013tables](#)). We calculate residuals as the difference between SSA’s “best guess” projection (intermediate-cost scenario / alternative II) and the historic statistics from SSA. Note that for 1981–1990, Trustees Reports have two intermediate-cost scenarios: alternative II-A and II-B. For these years, we follow subsequent Trustees Reports and use II-B as the “best guess” projection. To calculate the uncertainty interval, we use the minimum and maximum values of projected life expectancy across the three scenarios (four for 1981-1990).

Sources for SSA projections published in the annual Trustees Reports (TR) from 1978 to 2012:

Cost rate: Table 26 of TR 1978, Table 27 of TR 1980-1980, Table 28 of TR 1982, Table 29 of TR 1982-1983, Table 30 of TR 1984-1985, Table 28 of TR 1986, Table 26 of TR 1987-1991, Table II.F.13 of TR 1992-2000, Table IV.B1 of the supplemental single-year tables of TR 2001-2012.

Balance: For Trustees Reports from 1986, balance projections are available from the same tables as the cost rate projections. In 1978-1982 TR, the Trustees project the same scheduled tax rate across the cost scenarios (available in Table 25 for TR 1978, Table 26 for TR 1979-1980, Table 28 for TR 1981, and Table 29 for TR 1982). We subtract the cost rates across the scenarios from the scheduled income rate to obtain the range of balance projections. For 1983-1984 TR, the income tax rate varies slightly across the cost scenarios and is only published for the two intermediate projections (Table 27 for TR 1983 Table 28 for TR 1984). We are thus unable to evaluate coverage of SSA’s uncertainty intervals for projections made in these two reports.

Trust fund ratio: Table 28 of TR 1978, Table 29 of TR 1979-1980, Table 31 of TR 1981, Table 32 of TR 1982-1983, Table 33 of TR 1984-1985, Table 31 of TR 1986, Table 29 of TR 1987, Table

31 of TR 1988-1990, Table 32 of TR 1991, Table II.F.19 of TR 1992-1994, Table II.F.20 of TR 1995-2000, Table IV.B3 of the supplemental single-year tables of TR 2001-2012.

Unanticipated costs We calculate the unanticipated costs due to demography by estimating the number of unanticipated OASI beneficiaries for a given Trustees Report and projection year. In order to estimate the number of unanticipated beneficiaries, we determine the proportionate change in age-specific mortality rates that would correspond to the forecast error in life expectancy for each Trustees Report and projection year. For each age 65 years and older, we multiply the age-specific population count by the difference between the counterfactual age-specific mortality rate necessary to achieve the projected life expectancy and the observed age-specific mortality rate. This product represents the number of unanticipated beneficiaries for a given age and sex. We then sum across all ages for males and females to estimate the total number of unanticipated beneficiaries. For example, the 2005 Trustees Report projected year 2010 life expectancy at age 65 for males to be 16.6 years while observed life expectancy in 2010 equaled 17.9 years. This forecast error of 1.3 years corresponds to a 18.75% increase in observed age-specific mortality rates, or 150,844 unanticipated male beneficiaries.

To obtain unanticipated cost due to under-estimating life expectancy, we then multiply the number of unexpected OASI beneficiaries by the average annual benefits per OASI beneficiaries in constant 2010 dollars. We calculate average annual benefits per OASI beneficiary in current dollars by dividing total OASI expenditures ([j.mp/OASIexpend](#)) by total OASI beneficiaries ([j.mp/OASIbenif](#)). We convert average annual benefits from current dollars to constant 2010 dollars using the CPI-U series available from the Bureau of Labor Statistics ([bls.gov/data](#)).

We calculate total unanticipated costs as the difference between observed total OASDI expenditures in current dollars ([j.mp/OASDIexpend](#)) and projected OASDI expenditures. We convert unanticipated costs from current dollars to constant 2010 dollars using the CPI-U series available from the Bureau of Labor Statistics. Forecasts of total OASDI expenditures in billions of current dollars for 2000-2010 Trustees Reports are found in Table III.B3 of 2000 TR, Table VI.E9 of 2001-2002 TR, Table VI.F9 of 2003-2004 TR, and Table VI.F8 of 2005-2010 TR.

2 Main Figures in Color

Figures 1–7 in this Appendix are the color equivalents of Figures 1–7 from [Kashin, King and Soneji \(2015b\)](#).

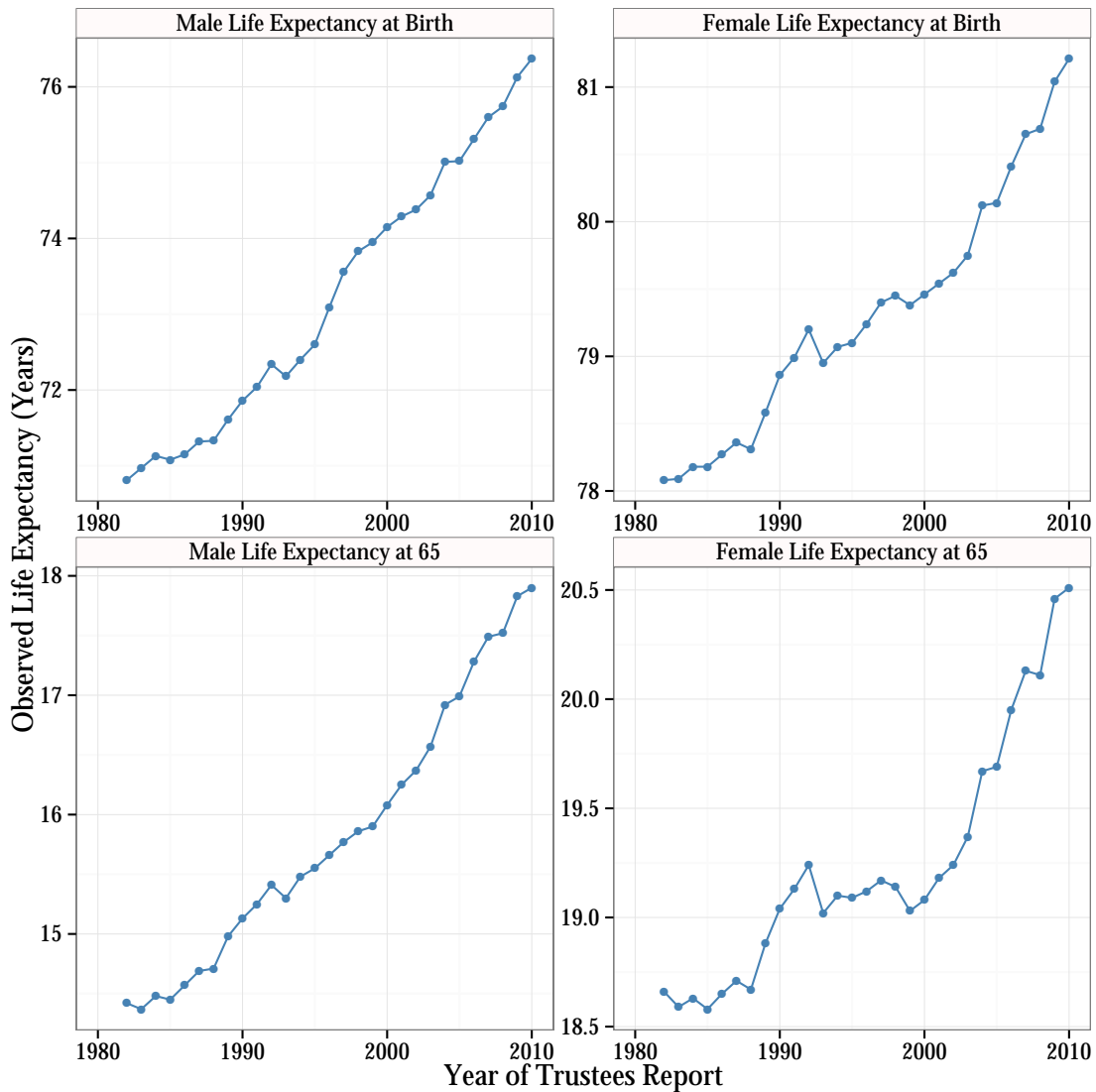
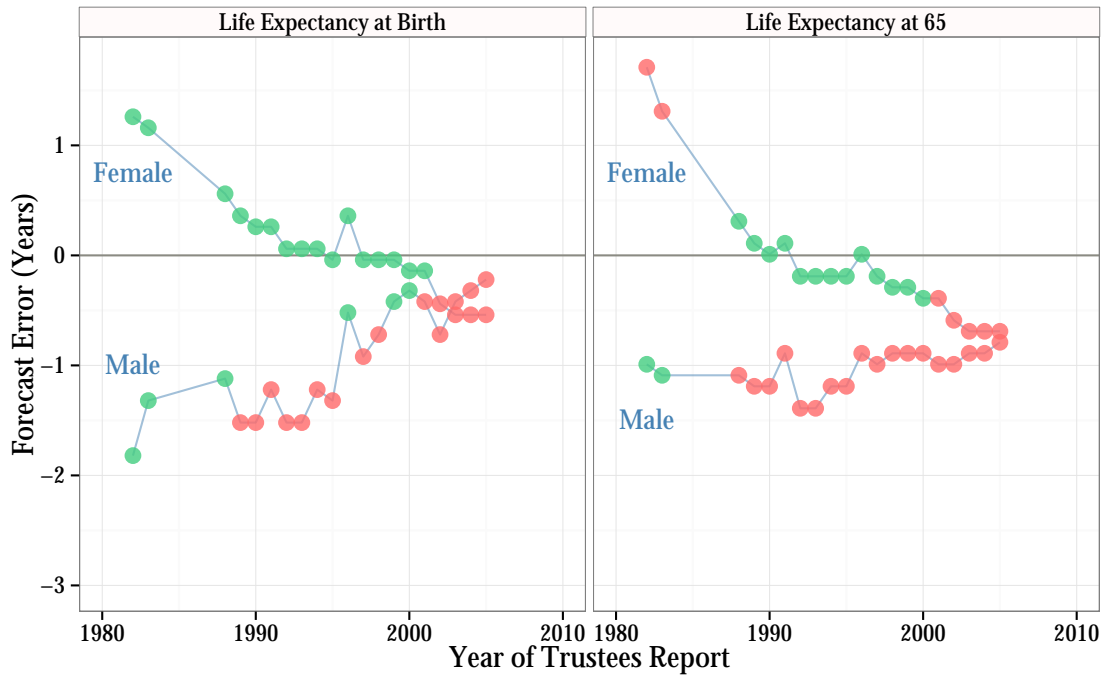


Figure 1: Observed Period Life Expectancy. As described in the text, “period life expectancy” for a year is a single-number summary of all the age-specific mortality rates for that same year and is interpreted as the average number of years a person could expect to live if he or she experienced the mortality rates of a given year over the course of their life.

(a) 2005 Life Expectancy Forecast Error



(b) 2010 Life Expectancy Forecast Error

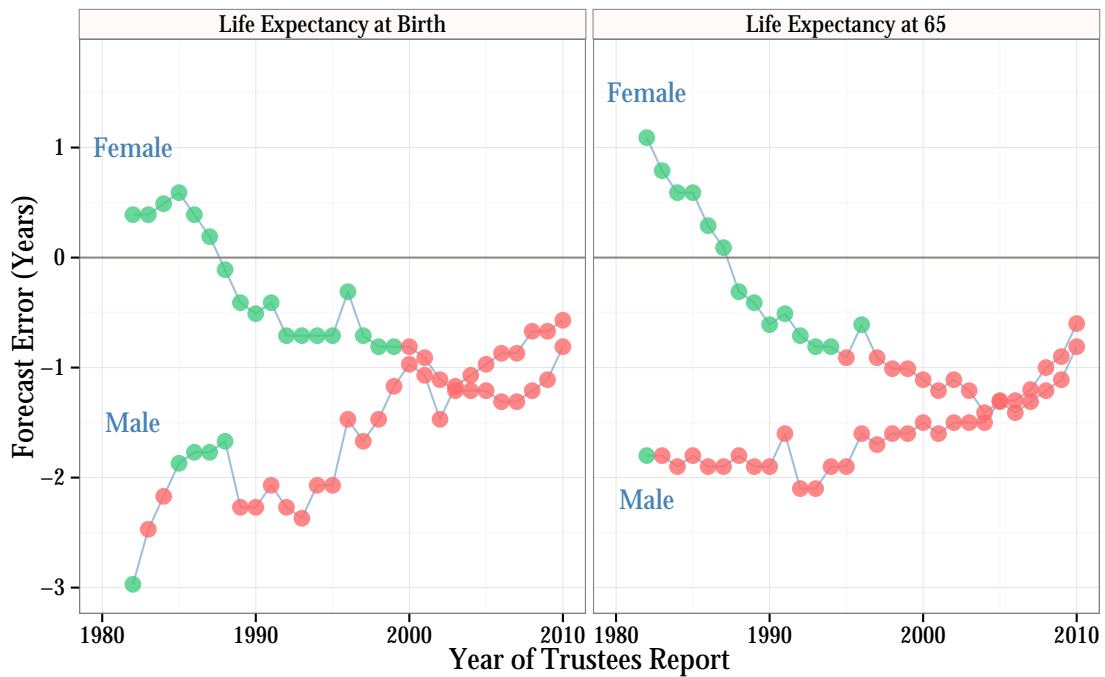


Figure 2: Forecast Error of Life Expectancy in 2005 (panel a) and 2010 (panel b) by Year of Trustees Report. Dots colored green when truth falls within SSA uncertainty intervals. Dots colored red when the truth falls outside SSA uncertainty intervals.

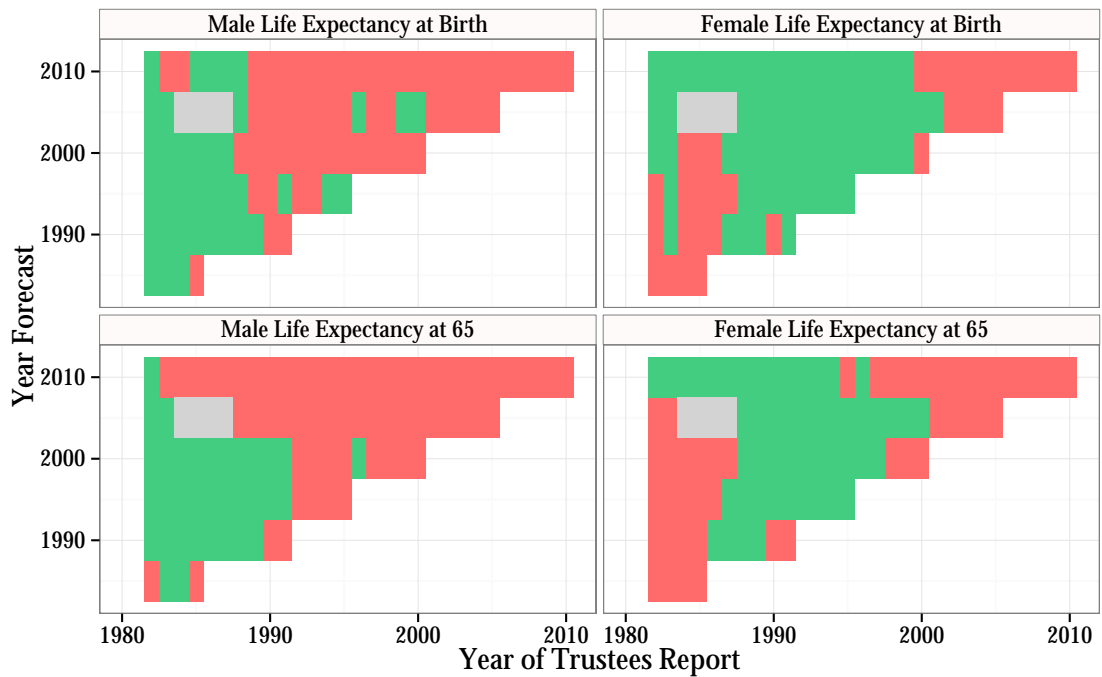


Figure 3: Uncertainty interval coverage by year of Trustees Report and year of forecast. Green indicates uncertainty interval covered the truth, red indicates that it did not, and gray indicates that SSA did not provide an interval. Contemporaneous forecast error is possible because of the time lag (typically three to four years) in finalizing mortality data.

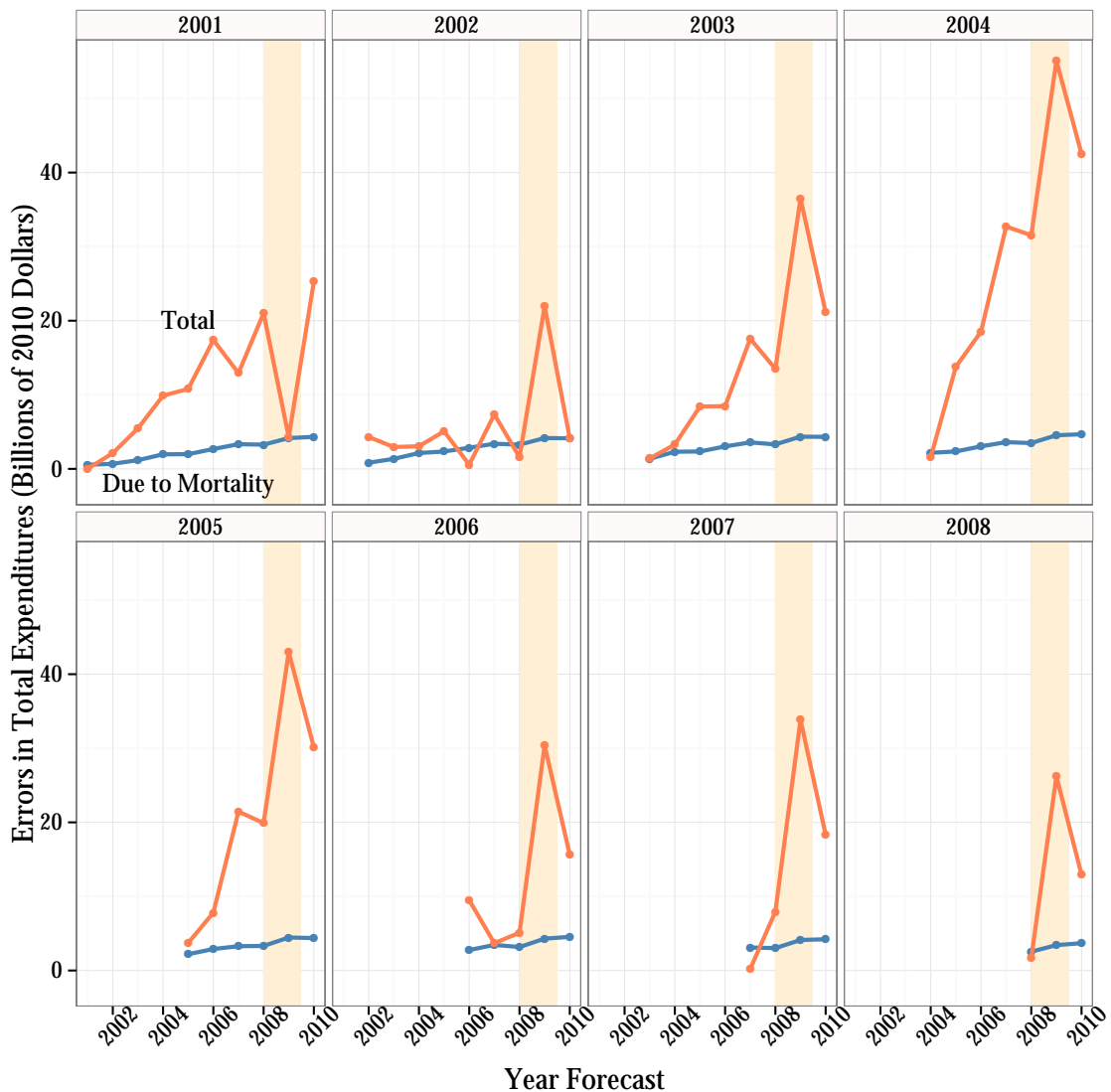


Figure 4: Cost of Mortality Forecasting Errors (in billions of 2010 dollars). Each panel of the figure corresponds to a Trustees Report. Within each panel, we plot the forecast error in total Social Security expenditures (dark orange lines) and the forecast error in total Social Security expenditures due to mortality forecasting errors (blue lines). Finally, we represent the Great Recession as a vertical light orange region.

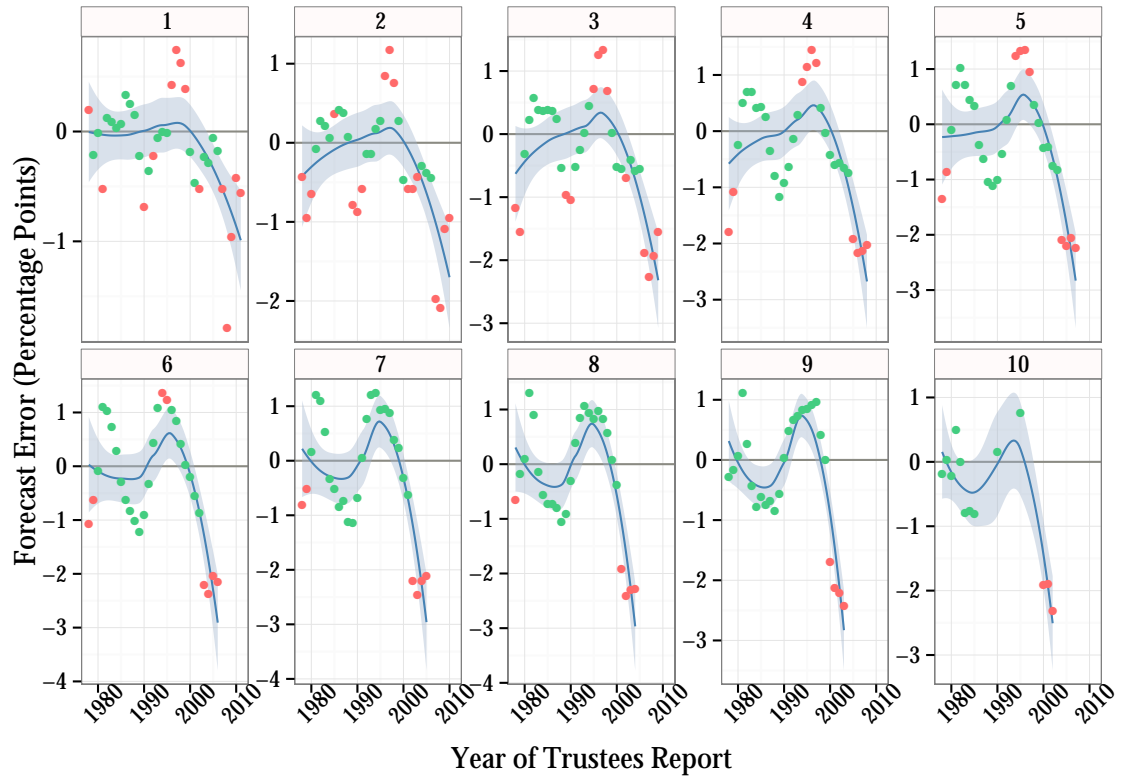


Figure 5: Cost Rate Forecasting Errors. Forecast errors in the cost rate (vertically) by the year of the forecast (horizontally) by how many years into the future is being forecast (in the title of each panel). Cost forecasts are overestimates if positive and underestimates if negative. Points are green if the error is within SSA's uncertainty interval and red otherwise.

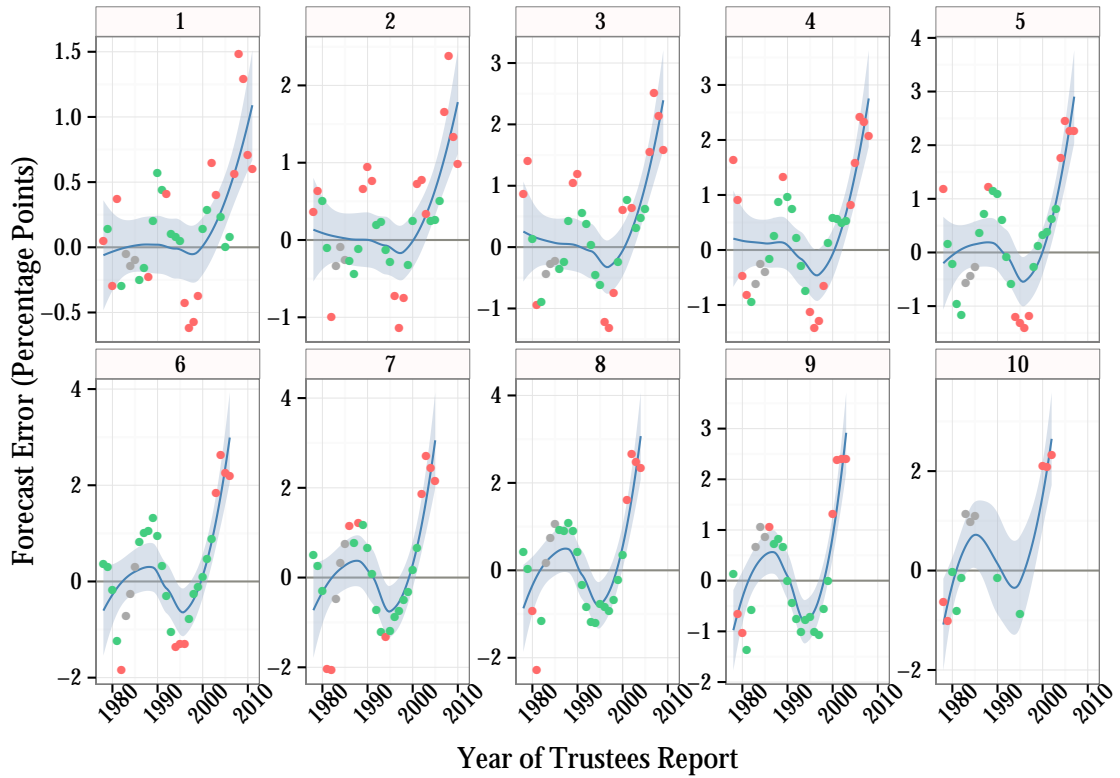


Figure 6: Balance Forecasting Errors. Forecast errors in balance (vertically) by the year of the forecast (horizontally) by how many years into the future is being forecast (in the title of each panel). Positive errors overestimate Trust Fund assets; negative errors underestimate them. Points are green if the error is within SSA's uncertainty interval and red otherwise.

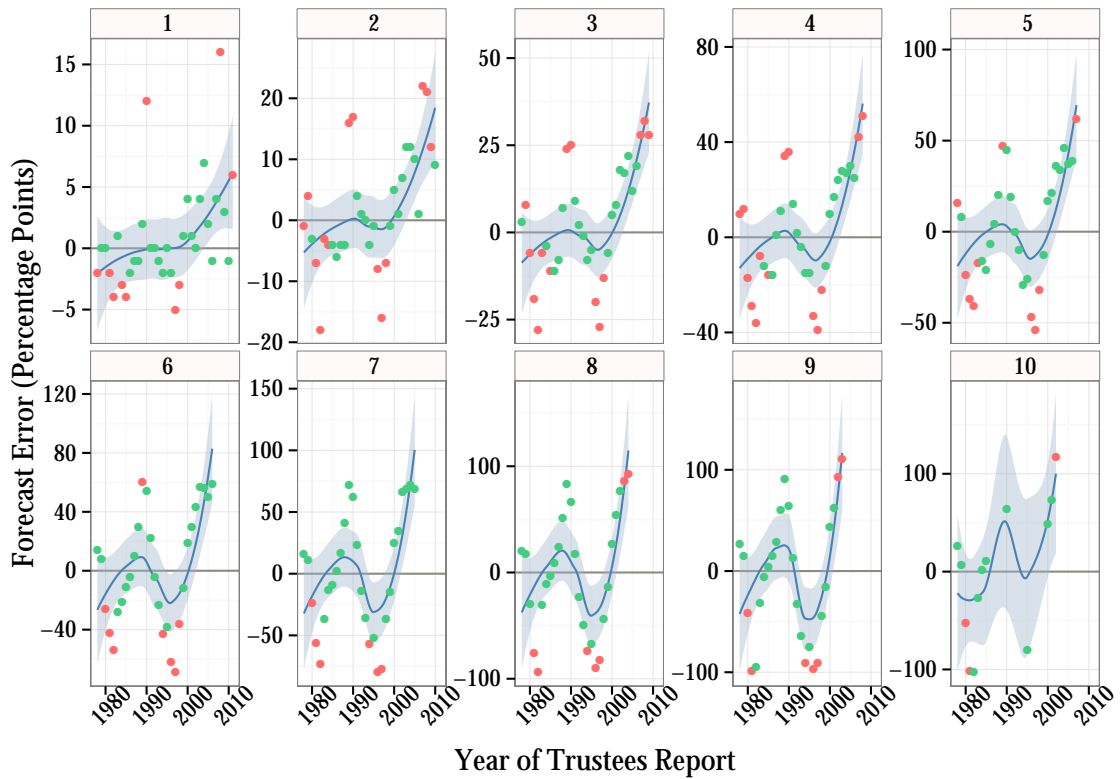


Figure 7: Trust Fund Ratio Forecasting Errors. Forecast errors in the trust fund ratio (vertically) by the year of the Trustees Report forecast (horizontally) by the number of years into the future being forecast (in the title of each panel). Positive errors overestimate Trust Fund assets; negative errors underestimate them. Points are green if the error is within SSA's uncertainty interval and red otherwise.

3 Demographic Results Using SSA Mortality Data

In Figures 8–10 in this Appendix, we replicate the analyses in Figures 1–3, respectively, from [Kashin, King and Soneji \(2015b\)](#) by replacing mortality data from the Human Mortality Database with that computed by SSA. Only small differences between the sources of data are revealed by this analysis.

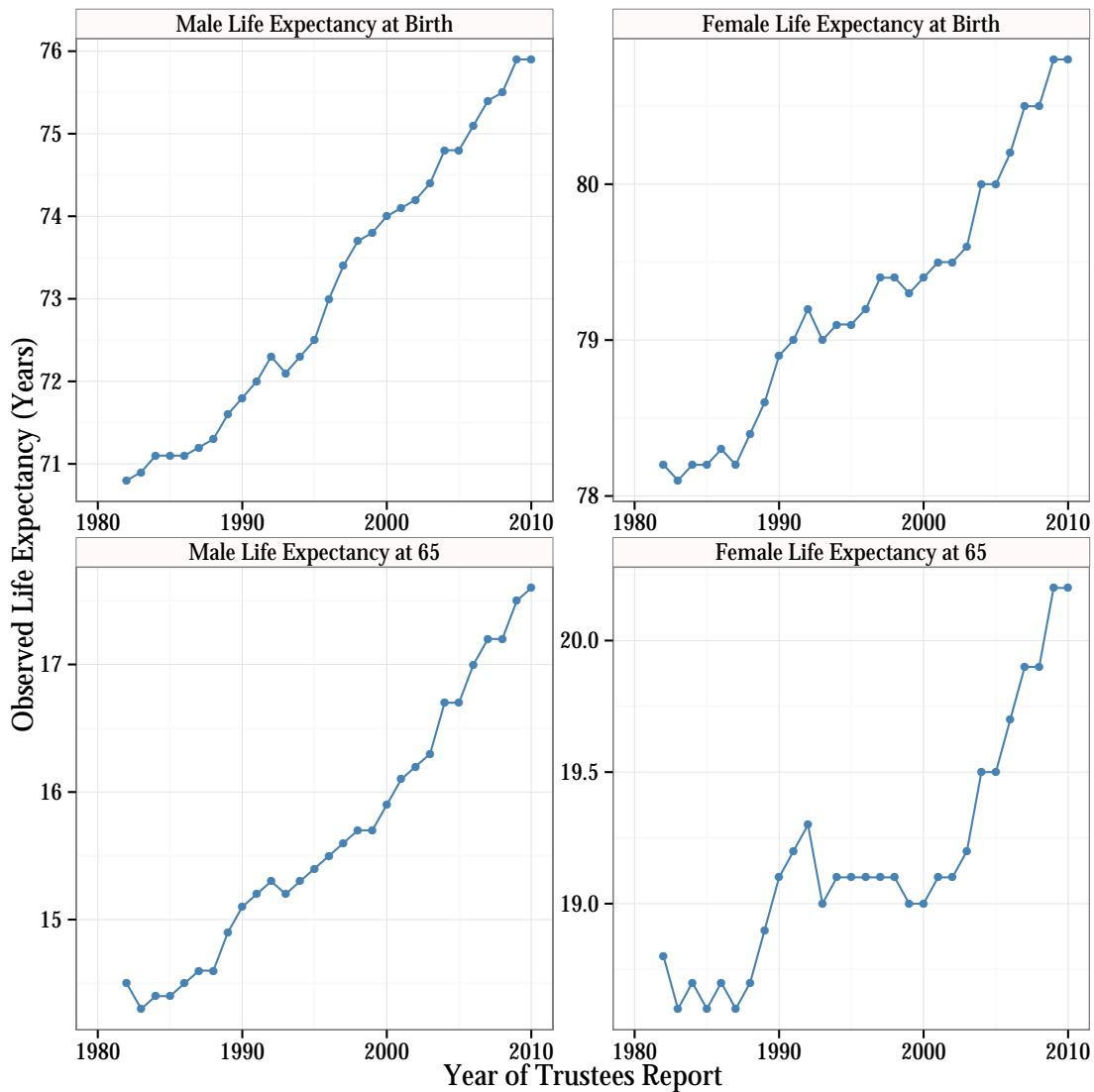
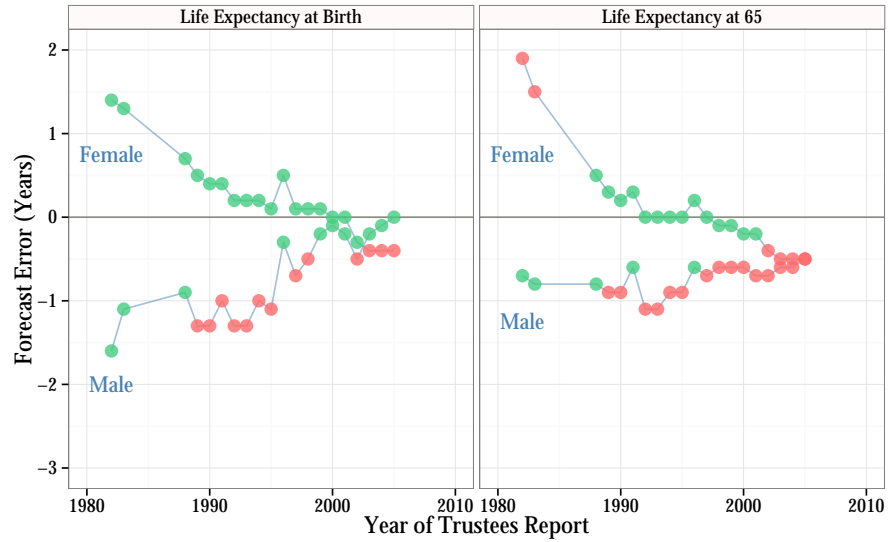


Figure 8: Observed Period Life Expectancy (Source: SSA). As described in the text, “period life expectancy” for a year is a single-number summary of all the age-specific mortality rates for that same year and is interpreted as the average number of years a person could expect to live if he or she experienced the mortality rates of a given year over the course of their life.

(a) 2005 Life Expectancy Forecast Error



(b) 2010 Life Expectancy Forecast Error

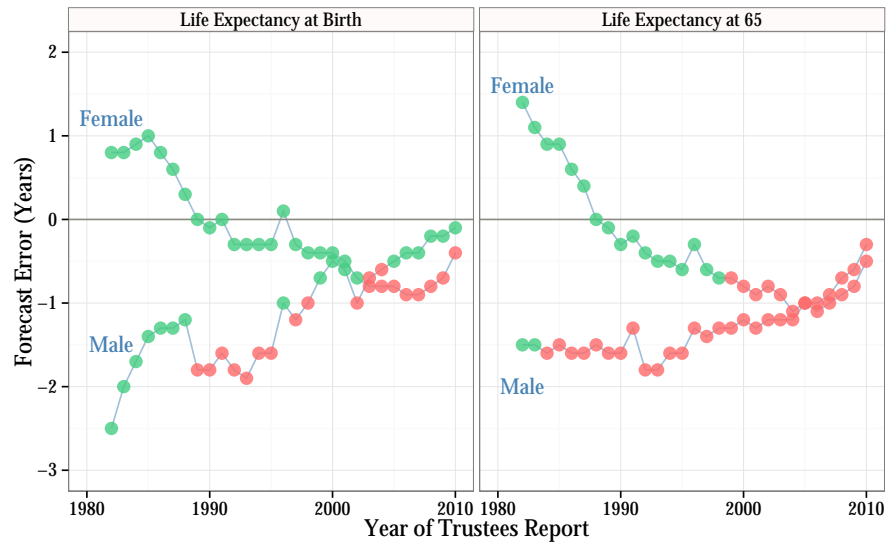


Figure 9: Forecast Error of Life Expectancy in 2005 (panel a) and 2010 (panel b) by Trustees Report, with data from SSA. Dots are colored green when truth falls within SSA uncertainty intervals. Dots are colored red when the truth falls outside SSA uncertainty intervals.

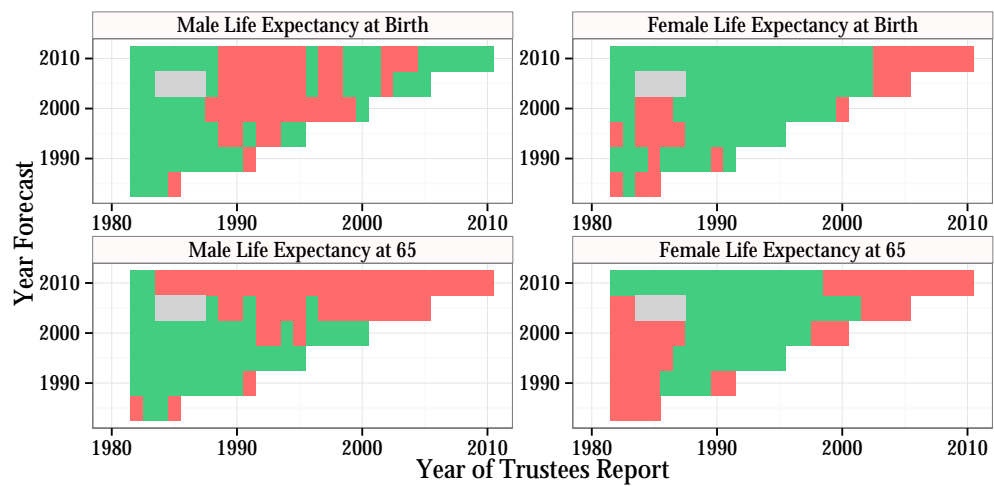


Figure 10: Uncertainty Interval Coverage, with data from SSA. Green indicates uncertainty interval covered the truth, red indicates that it did not, and gray indicates that SSA did not provide an interval.

4 Life Expectancy Forecast Errors for 1985–2010

We summarize in Figures 11 and 12 the forecast error for all available years of life expectancy (forecasts for years 1985–2010). The four graphs in this figure correspond to the four demographic variables from Figure 1. Again, each graph plots forecast error (vertically) by the year of the Trustees Report (horizontally). Each graph also includes a smoothed line (locally weighted scatterplot smoothing [LOESS]) weighted by proximity of the forecast (allowing for the fact that forecasting years farther into the future should be more difficult), along with a 95% confidence interval. The general pattern from the 2005 and 2010 demographic forecasts in Figure 2 can be seen here as well.

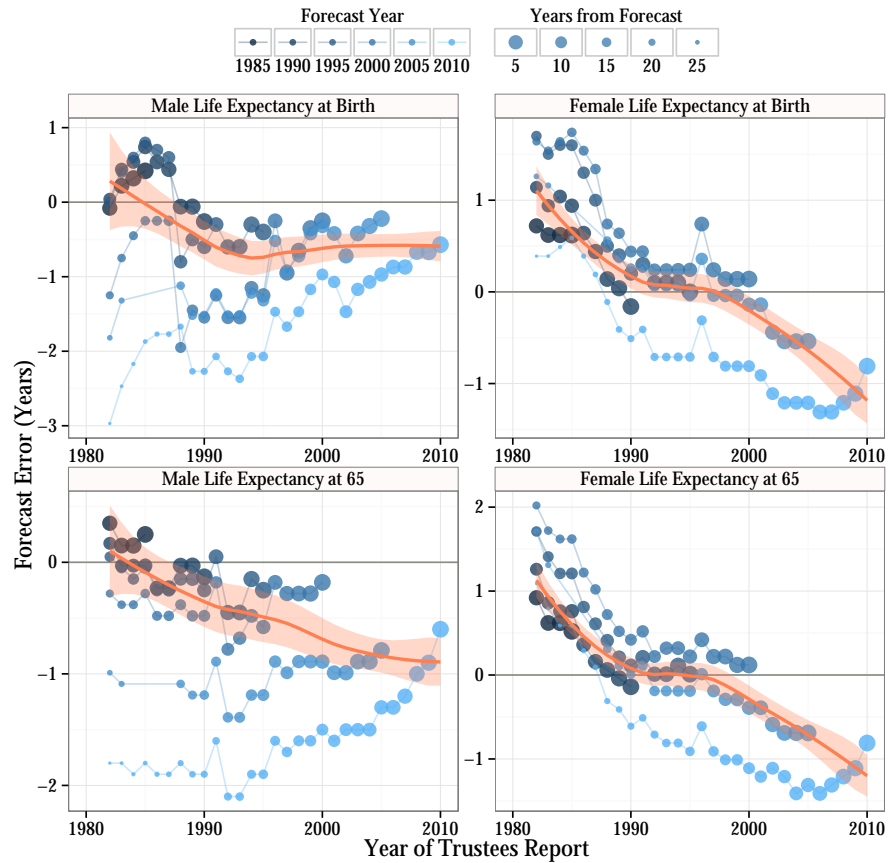


Figure 11: Life Expectancy Forecasts: Errors by Trustees Report Year by Forecast Year, with data from HMD. Larger dots indicate forecasts nearer to the date the forecast was made. Smoothed line weighted by proximity to year forecast.

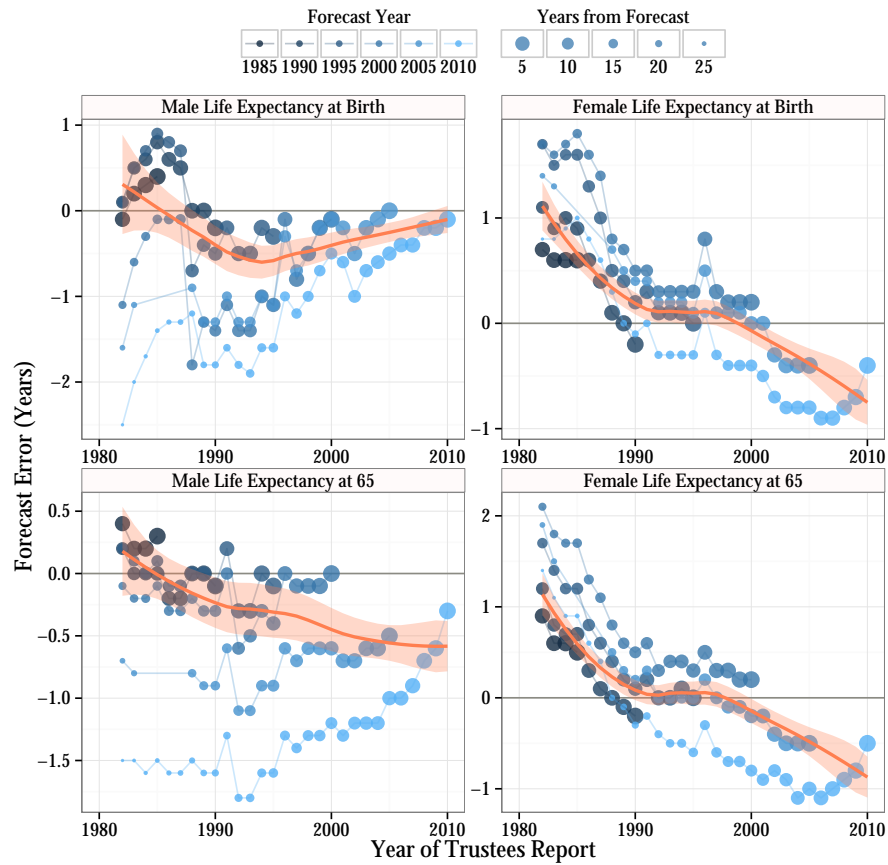


Figure 12: Life Expectancy Forecasts: Errors by Trustees Report Year by Forecast Year, with data from SSA. Larger dots indicate forecasts nearer to the date the forecast was made. Smoothed line weighted by proximity to year forecast.

5 Total Fertility Rate Forecast Errors

Our results for the analysis of total fertility rate forecasts appear in Figure 13 for data from the Human Mortality Database and Figure 14 for SSA data. The two figures are quite similar. Results in both are relatively unremarkable, except for consistent overconfidence nearer to the date forecast. See for example the bottom right graph in both figures which has disproportionate numbers of red squares near the diagonal on the graph (where the year of Trustees Report is close to the year forecast).

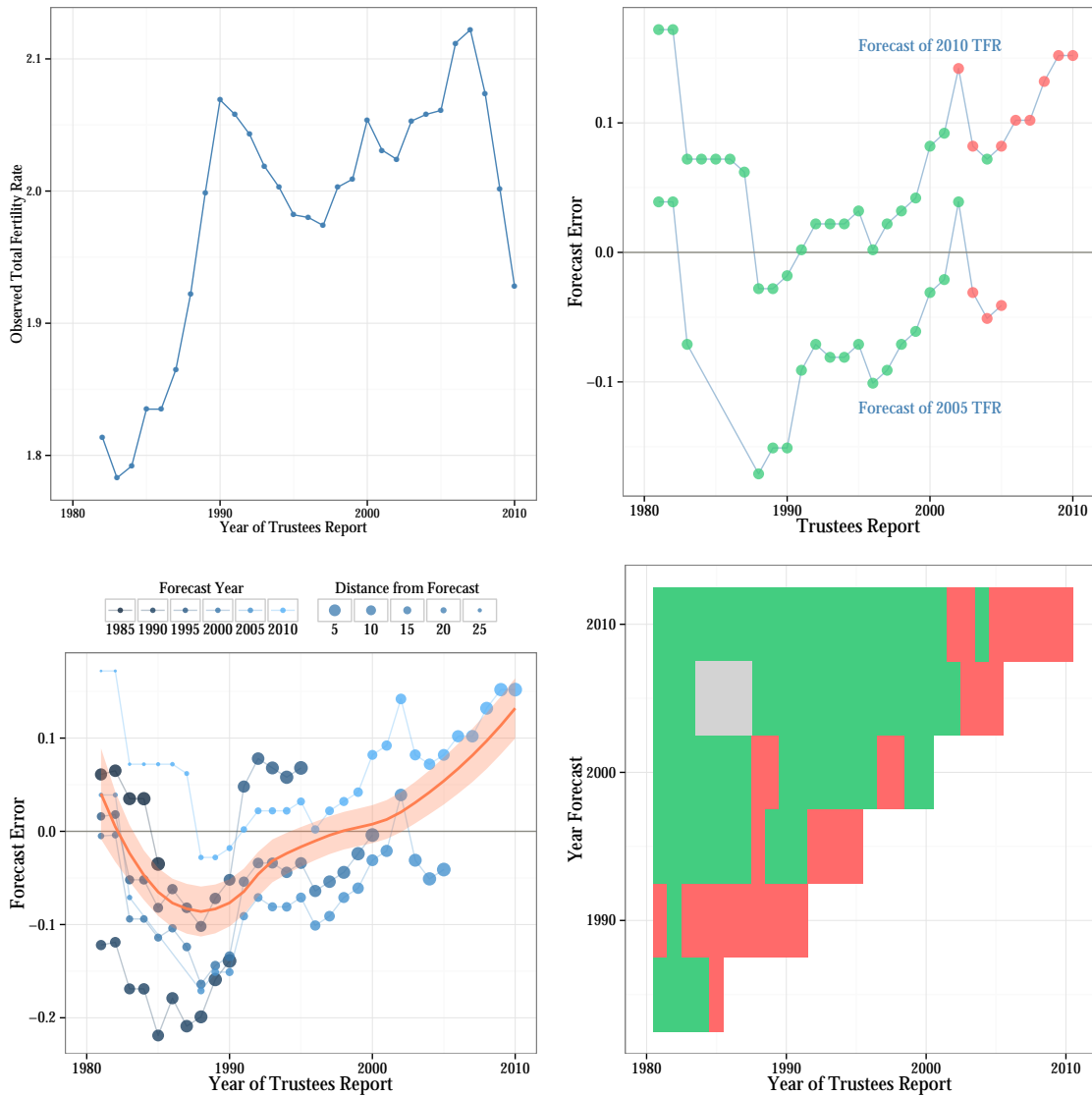


Figure 13: Clockwise from top left: a) Observed total fertility rate from Human Fertility Database; b) Forecast error of total fertility rate for 2005 and 2010 by Trustees Report. Dots are colored green when truth falls within SSA uncertainty intervals. Dots are colored red when the truth falls outside SSA uncertainty intervals; c) Errors by trustees report year by forecast year. Larger dots indicate forecasts nearer to the date the forecast was made. Smoothed line weighted by proximity to year forecast; d) Uncertainty interval coverage. Green indicates uncertainty interval covered the truth and red indicates that it did not.

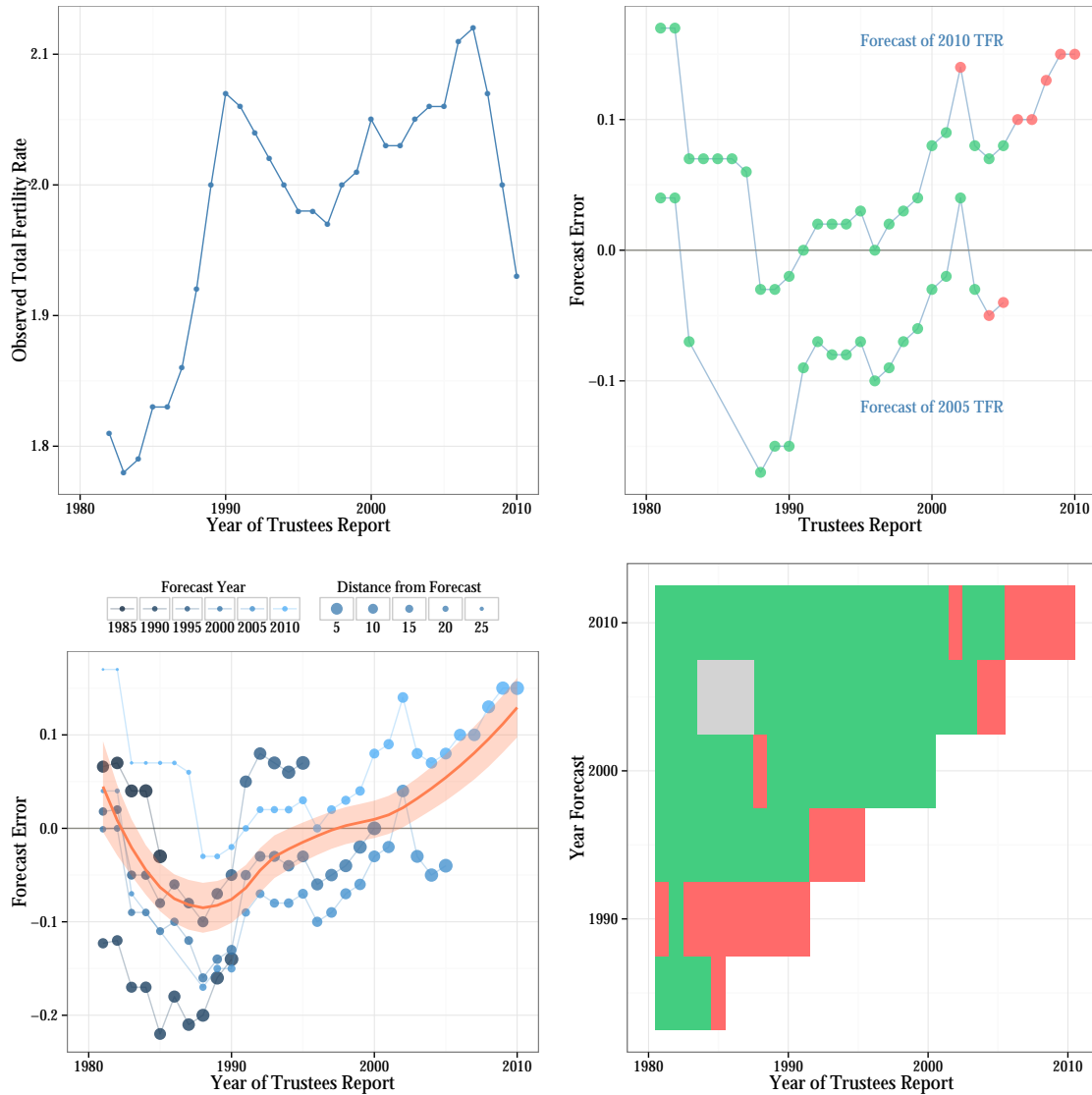


Figure 14: This figure replicates the analyses in Figure 13 by replacing fertility data from the Human Mortality Database with that from SSA.

6 Migration Forecast Errors

Our results for the analysis of net legal immigration forecasts appear in Figure 15. For a brief discussion of this figure, see “A Note About Fertility & Immigration” in Section 1 of [Kashin, King and Soneji \(2015b\)](#).



Figure 15: Forecast Error of Legal Migration in 2005 (panel a) and 2010 (panel b) by Year of Trustees Report. Dots colored green when truth falls within SSA uncertainty intervals. Dots colored red when the truth falls outside SSA uncertainty intervals.

References

- Kashin, Konstantin, Gary King and Samir Soneji. 2015a. “Replication Data for: Systematic Bias and Nontransparency in U.S. Social Security Administration Forecasts.” UNF:5:1oerGFXQ0Bu9bcMFU5/t2A== <http://dx.doi.org/10.7910/DVN/28122> Harvard Dataverse [Distributor] V1 [Version]. 2
- Kashin, Konstantin, Gary King and Samir Soneji. 2015b. “Systematic Bias and Nontransparency in U.S. Social Security Administration Forecasts.” *Journal of Economic Perspectives* . In press. 2, 4, 11, 18