International Statistical Review (2004), **72,** 1, 93–106, Printed in The Netherlands

Simpler Probabilistic Population Forecasts: Making Scenarios Work

Joshua R. Goldstein

Office of Population Research, Princeton University, Wallace Hall, Princeton, NJ, USA. E-mail: josh@princeton.edu

Summary

The traditional high-low-medium scenario approach to quantifying uncertainty in population forecasts has been criticized as lacking probabilistic meaning and consistency. This paper shows, under certain assumptions, how appropriately calibrated scenarios can be used to approximate the uncertainty intervals on future population size and age structure obtained with fully stochastic forecasts. As many forecasting organizations already produce scenarios and because dealing with them is familiar territory, the methods presented here offer an attractive intermediate position between probabilistically inconsistent scenario analysis and fully stochastic forecasts.

Key words: Age structure; Population forecasting; Population size; Scenarios; Stochastic; Uncertainty.

1 Introduction

Traditionally, population forecasts have included high, medium, and low variants, with the medium being the most likely outcome. The high and low scenarios are said to be "plausible" but lack any probabilistic interpretation. High and low scenarios tend to produce extremes in either population size or in population age structure, but not in both at once. Scenarios typically have smooth time trajectories, assigning zero probability to fluctuations in fertility and mortality that occur in real populations. A recent evaluation concludes that "no consistent probabilistic interpretation *can* be given to the high-low scenarios in population forecasts". (National Research Council, 2000).

Stochastic population forecasts have been developed that try to address these difficulties. A number of approaches have been tried. Approaches based on stochastic renewal (Lee & Tuljapurkar, 1994; Keilman, Pham & Hetland, 2002; Alho & Spencer, 1991) have used time series analysis to forecast future mortality and fertility rates which are allowed to fluctuate randomly over time. (Typically, migration remains deterministic.) Monte Carlo methods are then used to produce a large number of future population trajectories. The quantiles of these simulated trajectories give the prediction intervals for measurements of population size, and age-structure. An analytic theory of stochastic renewal has also been developed (Cohen, 1986; Lee & Tuljapurkar, 1994; Alho & Spencer, 1991), but modern computing power makes the Monte Carlo approach simpler in most cases. The world forecasts by Lutz, Sanderson & Scherbov (2001) have used a hybrid of methods including expert opinion, scenarios, and auto-correlated time series to produce probabilistic projections.

Fully stochastic forecasts quantify the uncertainty in population size, age groups, and age structure in a probabilistically consistent way. This internal consistency allows the incorporation of mortality and fertility uncertainty, simultaneous prediction intervals for several age groups and functions of age groups, and population outcomes to be integrated over time. This last feature is of particular advantage when assessing the viability of national pension systems for which the age structure of a population over many decades is of interest (Lee & Tuljapurkar, 1998).

Despite these advantages, population projectors have been slow to switch over to fully stochastic forecasts. Their reluctance is not surprising. Stochastic forecasts are data-intensive, requiring detailed historical information to estimate the variability of vital rates. They also assume that variability in the future will be similar to that of the past. Finally, from a technical point of view, stochastic forecasts are difficult to produce—whereas scenario projections require relatively little training and can be produced with existing software, stochastic forecasts have required research scientists knowledgeable about time series estimation and customized programming.

Moreover, scenarios are easy to explain to users and easy for users to interpret. The middle scenario is interpreted as the most likely outcome (Keyfitz, 1972). “High” and “low” may lack assigned probabilities but do have simple conditional interpretations. *If* fertility and mortality follow the hypothesized paths *then* the impact for future population will be as shown. This if-then approach translates vital rate trajectories into population outcomes, a rough sensitivity analysis of how population size and age-structure depend on levels of fertility, mortality, and migration

This paper addresses the reluctance of forecasters to make the leap from using scenarios to fully stochastic forecasts. We do this by describing the conditions under which the results of scenario-based forecasts closely approximate those that could have been obtained using a fully stochastic approach.

The procedure presented here is a compromise. It does not produce exactly correct results, but it yields surprisingly accurate figures with little additional effort on the part of the forecasters. Those who use this approach have the benefits of both using a methodology with which they are familiar and obtaining the desirable features of probabilistically consistent uncertainty intervals. To do this, we first give scenarios a specific probabilistic interpretation; then we combine the results for different age-segments in a probabilistically sensitive manner.

Past comparisons between scenarios and stochastic forecasts have compared forecasts that differ in many dimensions including both the expected level of fertility and mortality and the range of fertility and mortality (Lee & Tuljapurkar, 1994; Keilman *et al.*, 2002). The approach here is to calibrate the two approaches so that a comparison can reveal the differences and similarities intrinsic to the two methods. Lutz & Scherbov (1998) used a similar approach but a different method of calibration.

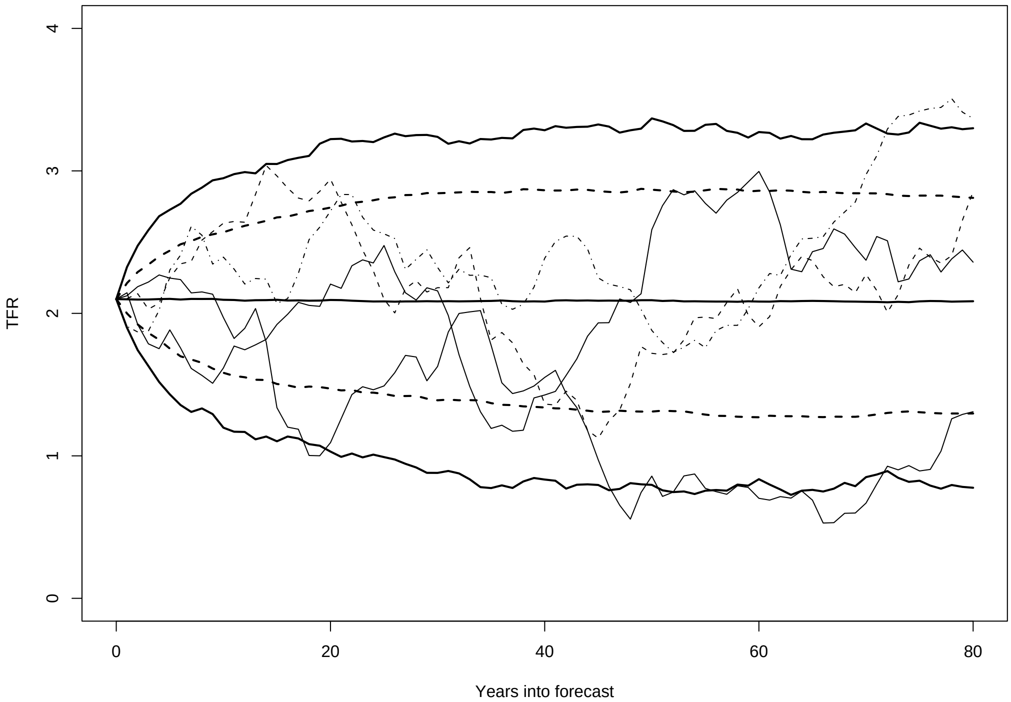
The paper begins with a discussion of how to give of scenarios, particularly fertility scenarios, a specific probabilistic interpretation. Next we consider how to combine the results of forecasts for different age-segments in a probabilistically sensitive manner. Throughout, we compare the results of this modified scenario approach with published fully stochastic forecasts by Lee & Tuljapurkar (1994).

2 The Modified Scenario Approach

Our first task is to assign a probabilistic interpretation to the vital rates scenarios. Figures 1 and 2 show some sample paths from stochastic forecasts of fertility and life expectancy. The distance between the bold solid lines gives the 95 percent uncertainty intervals for fertility and life expectancy in any given year. The vertical span between the heavy dashed lines in the fertility figure gives the 95 percent interval of the long-term average of fertility, which will be described below. The intervals for for the average are narrower than the intervals for a given year because because the ups and downs in fertility cancels out to some extent.

A key difference between scenario-based forecasts and stochastic forecasts is that the scenarios incorporate nearly perfect temporal autocorrelation. If, for example, fertility is high in the first 5 years of the scenario forecast, it is assumed to remain high for the entire forecast. Stochastic forecasts, on the other hand, use time series models to incorporate an empirically estimated amount

Figure 1. *Sample paths and forecast intervals of stochastic forecast of total fertility rate.*



Stochastic fertility model is a mean-constrained ARMA(1,1) model with coefficients given in Lee & Tuljapurkar (1994). The figure shows 4 sample paths generated with Monte Carlo simulation; these can be identified by their large ups and downs. The heavy solid U-shaped outer bounds give the 95 percent prediction interval for the TFR in any year, with the heavy middle line giving the median TFR. The heavy dashed U-shaped inner bounds give the 95 percent prediction interval for the cumulative average of the TFR up to that date. Both prediction intervals were estimated from the quantiles of 1000 simulated runs. Note that the bounds on the cumulative average cover roughly the same interval from years 20 through 80.

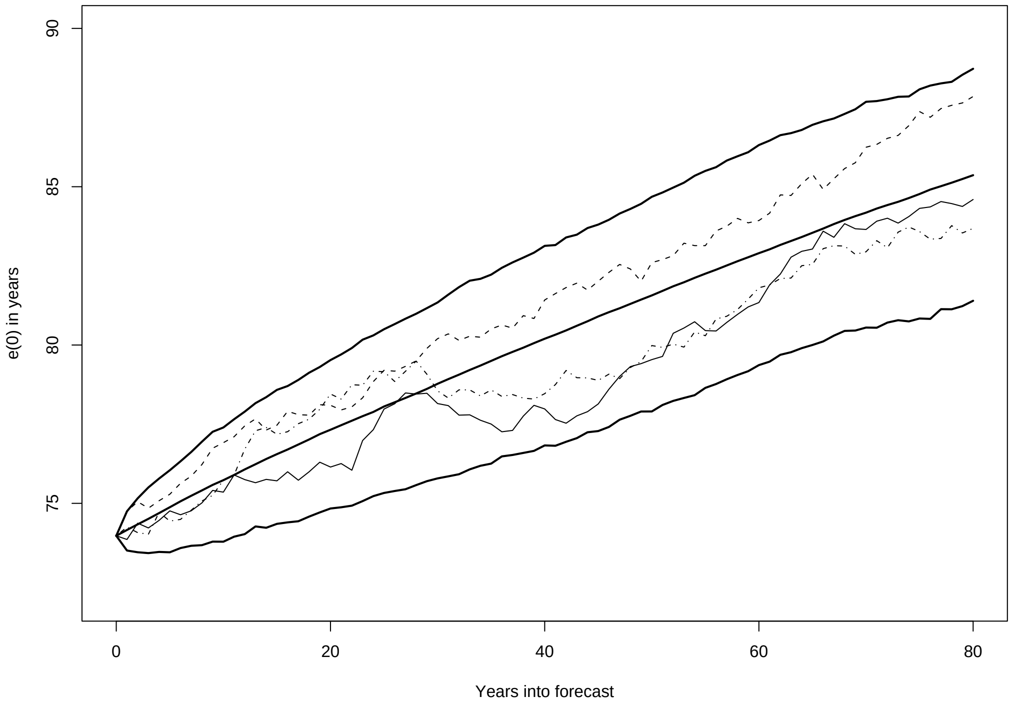
of autocorrelation over time. In the illustration, the fertility model is a constrained mean ARMA(1,1) model (Lee, 1993) and the mortality model is a random walk with drift (Lee & Carter, 1992a).

In an earlier paper, Lutz & Scherbov (1998) used the prediction interval on fertility in a given year to construct scenarios. They found that this produced larger estimates of uncertainty than the fully stochastic approach because—as Lee (1999) points out the stochastic forecasts incorporate some "cancellation of errors", whereby a path that has high fertility in one year is likely not to be the same path that has high fertility in the next year.

An alternative approach, adopted here, is to calibrate scenarios to a smoothed version of the stochastic realizations, averaging out the highs and lows. Specifically, we construct scenarios based on the uncertainty in the cumulative average of fertility from the beginning of the forecast to the date of interest. The cumulative average $\bar{X}(t)$ is defined as $\sum_0^t X(i)/t$.

An interesting feature of the cumulative average is that its variance is nearly constant after about year 20 of the forecast. In the absence of serial correlation, the variance of the cumulative average would shrink over time. However, the observed temporal autocorrelation is sufficient to keep the uncertainty interval nearly stable from year 20 through year 80 of the forecast, and shrink significantly

Figure 2. *Sample paths and forecast interval of stochastic forecast of life expectancy.*



Stochastic mortality model is a random walk with drift (Lee & Carter, 1992a) with specification and coefficients given by Lee & Tuljapurkar (1994); life expectancy is not forecasted directly but rather is a function of the mortality index. The figure shows 3 sample paths generated with Monte Carlo simulation; these can be identified by their large ups and downs. The heavy solid U-shaped outer bounds give the 95 percent prediction interval for life expectancy in any year, with the heavy middle line giving the median life expectancy, as estimated from the quantiles of 1000 simulated runs.

only later on. This property is convenient because it means that the bounds on the calibrated fertility scenarios can be approximated by a moving moving average with a window of several decades, and that the choice of window width will have a small effect on the interval coverage. The stability of the uncertainty interval of the cumulative average is a robust finding. It also is produced by other autoregressive specifications (e.g., AR(1)) and in other populations in the industrialized world (e.g. France). However, it is possible to construct models that have less stable uncertainty intervals for the cumulative average of forecast fertility (Congressional Budget Office, 2001).

For fertility the use of the variation in long-term averages seems justifiable on several counts. First, forecasters are generally not interested in the number of births in a particular year, but rather in the number of births over broad time intervals that determine the size of future broad age groups. Second, the wide span of reproductive years for any cohort means that over the course of generations, the number of potential parents in a broad age interval rather than the numbers of potential parents in a single year of age are what—in combination with period fertility—determine the number of births in a given year.

Analytically, it is difficult to show that the cumulative average is the right kind of average to use in age-structured populations experiencing stochastic renewal. However, a simple non-overlapping generation model in which one generation reproduces the next can be used to show that calibrating scenarios to the cumulative average will give approximately the same variance as the fully stochastic approach. This is demonstrated in appendix B.

Stochastic forecasters have used the cumulative average to compare the variability of their fertility forecasts with scenarios (Lee, 1993; Congressional Budget Office, 2001). The modified scenario approach set out here simply takes this comparison one step further, interpreting the prediction interval bounds of the cumulative average as high and low scenarios.

The second step in improving scenario-based forecasts is to combine the uncertainty of different age groups in a probabilistically sensible manner. The temporal correlation of the scenarios for fertility assure that a large number of births in the first years of a forecast will always be accompanied by a larger than expected number of births later on. Similarly, high fertility and low mortality will create larger than expected population sizes across all ages, while low fertility and high mortality will do the opposite. The resulting estimates of variability in age structure will be distorted because of the perfect correlations that are produced using the scenarios. A way to correct this distortion is to take account of the magnitude of correlation that exists between broad age-segments.

Table 1

Age segment correlations and coefficients of variation from stochastic forecasts.

Forecast	Correlation			Coef. of Variation		
Year	O, W	Y, W	O, Y	Y	W	O
10	1.00	0.10	0.09	0.06	0.00	0.01
20	1.00	0.06	0.05	0.15	0.00	0.02
30	0.23	0.63	0.02	0.22	0.03	0.04
40	0.12	0.78	0.01	0.30	0.07	0.06
50	0.08	0.82	-0.01	0.39	0.12	0.08
60	0.04	0.83	-0.04	0.46	0.20	0.09
70	0.12	0.86	0.00	0.52	0.28	0.10
80	0.47	0.88	0.31	0.57	0.36	0.16

Estimated from simulation of stochastic forecasts as given in Lee & Tuljapurkar (1994) with 400 independent runs. Y is youth population aged 0–19, W is working-age population aged 20–64, and O is old-age population aged 65+. Coeficient of variation is ratio of the observed standard deviation to the observed mean.

Table 1 shows the correlations estimated between broad age groups over time for the fully stochastic forecast developed by Lee & Tuljapurkar (1994). The correlations between the working age and old age segments are close to zero for years 30 through 70, due to the independence of fertility and mortality assumed in the stochastic forecasts. The high correlations between the working and old age segments in the first two decades is due to the mortality model being used which assumes correlated effects across all ages in a given period (Lee & Carter, 1992a). This high degree of correlation can be neglected since there is virtually no variation in the size of the working age population. The working age and elderly age segments become positively correlated at the very end of the forecast when fertility uncertainty reaches the oldest age segment. Generational reproduction then introduces a positive correlation between the size of the elderly generation and those born afterwards. With these two caveats, there is little correlation between the size of the working and old-age population segments.

Generational reproduction also assures a high degree of correlation between generations for working and young age segments for forecast years 30 through 80. Before about year 30, the slight uncertainty in the number of working age is determined by mortality. After about year 30, the size of the first wave of births determines in large part the size of the second wave, producing the high correlations we see in Table 1. Although uncertainty in fertility rates also has an effect, it is of second order, particularly when there is positive autocorrelation of fertility rates over time.

The set of correlations and seen in Table 1 are based on my replication of the Lee and Tuljapurkar stochastic forecasts of the U.S. population. In general, a forecaster will not have a fully stochastic forecast in hand—that is why he or she is interested in using scenarios. Accordingly, rather than using the correlations produced by the U.S. forecasts, the approach we use to approximate the uncertainty in dependency ratios assumes independence ($\rho = 0$) between the elderly and the rest of the population and perfect correlation ($\rho = 1$) between the size of the young and working population.

Assuming zero correlation between the elderly and the rest of the population also allows one to treat fertility and mortality as independent factors for the time period we are considering, since the number of elderly is unaffected by future fertility for the first 60 or 70 years of the forecast.

3 Results

We now investigate how close the modified scenario approach comes to reproducing the uncertainty estimates from fully stochastic forecasts, assumed to give the true uncertainty of future population estimates. It should be kept in mind, however, that even fully stochastic forecasts omit many sources of uncertainty including estimation of the baseline population, specification error in time series models, and the probability of significant events that did not occur in the past, such as nuclear war (Lee & Carter, 1992b).

The stochastic forecasts of the United States population by Lee & Tuljapurkar (1994) constitute our benchmark case. (For details see appendix A.) For simplicity, we use the estimated Lee and Tuljapurkar time series models for fertility and mortality but omit migration and project only the female population.

Fertility scenarios are calibrated to the same 95 percent prediction bounds of *cumulative average* fertility produced by the stochastic forecasts—the heavy dashed lines in Figure 1. The life expectancy scenarios are calibrated to the 95 percent prediction interval of the forecast shown in Figure 2. Using this broader interval for life expectancy, rather than some kind of cumulative average, turns out to produce reasonable uncertainty ranges.

The observed 1990 U.S. female population is then projected according to three variants, which we label “large”, “small”, and “medium”. The large variant is for high fertility and low mortality. The medium variant is for medium fertility and mortality. The small variant is for low fertility and high mortality.

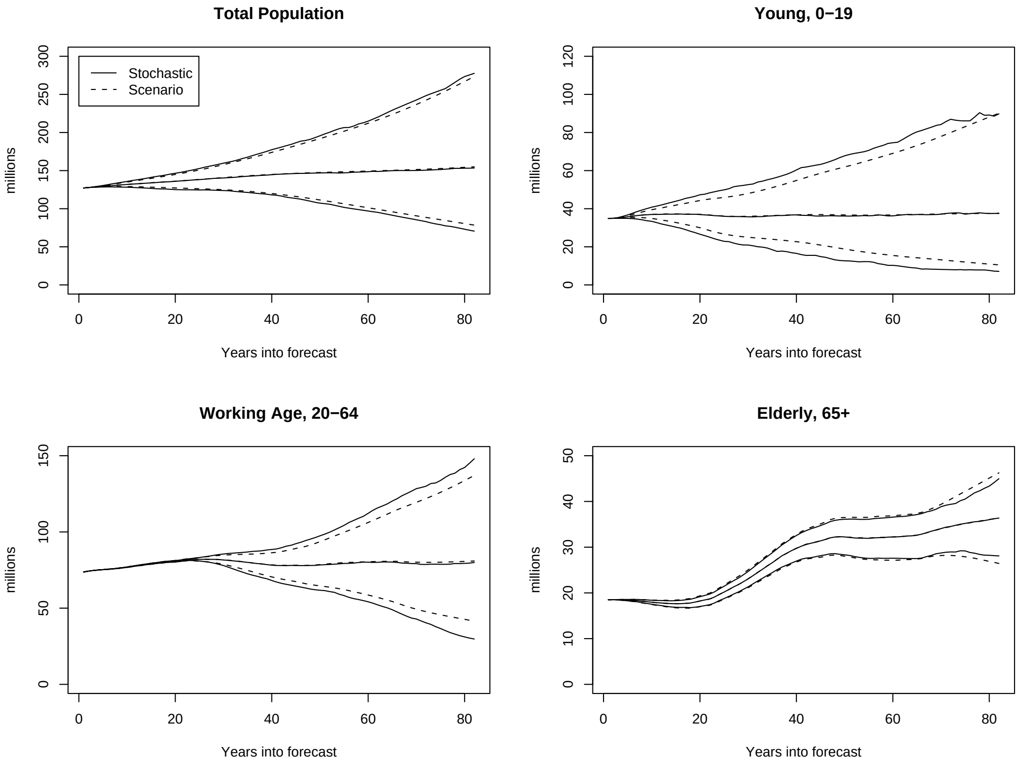
This choice of variants assures the greatest variation in age-segment sizes. As will be seen, a further round of calculation is needed to take into account the proper level of uncertainty on age structure.

3.1 Uncertainty in Total Population Size and the Size of Broad Age Segments

Figure 3 presents the total population size and the number of young, working and elderly estimated over time by the three variants. Fertility is the overwhelming source of uncertainty in population size (Alho, 1992), and we can see that the calibration of fertility scenarios to the prediction interval on the cumulative average of stochastic fertility has indeed created a very close match between the uncertainty intervals from the two approaches. This is a new finding. Past comparisons have shown that stochastic forecasts and scenarios have different ranges of uncertainty around age-segment sizes and total population (Lee & Tuljapurkar, 1994; Lutz & Scherbov, 1998). But no earlier comparisons have calibrated fertility so that its effective variation is similar in the two approaches.

The extremely close match for total population size between the two approaches results from two offsetting differences between the stochastic and scenario forecasts. The scenarios couple high

Figure 3. *Population size according to stochastic and scenario forecasts of U.S. females from 1990.*



In the scenario approach uncertainty for component age groups sums to the uncertainty in the population. In the stochastic approach, the uncertainty in the total population is less than the sum of the uncertainty of the components because of offsetting variability. Scenarios are calibrated to the uncertainty ranges in mortality and cumulative fertility.

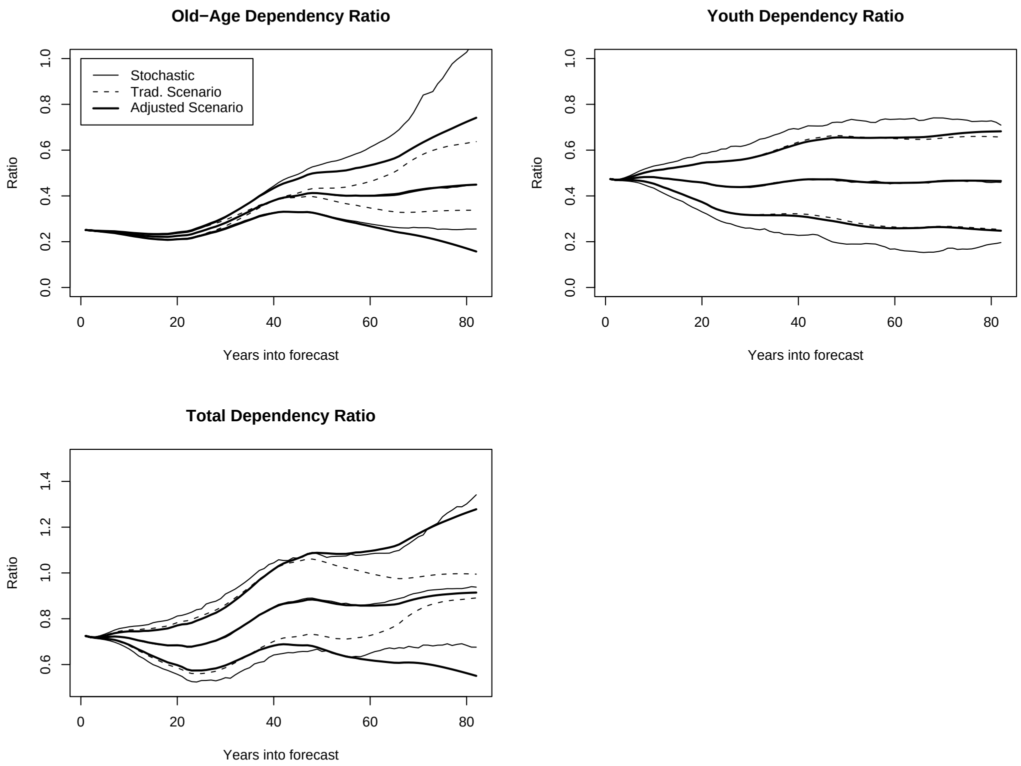
fertility with low mortality and vice-versa, inflating slightly the bounds on total population size compared to the independence between factors assumed in the stochastic forecast. On the other hand, we can see that the scenarios underestimate slightly the width of the interval covering each of the age segments, presumably because they do not allow the same degree of short-term variation in fertility rates that produces extremes in age-group sizes. The net result is that the scenarios produce almost exactly the same uncertainty ranges on total population size as the fully stochastic forecasts. An important qualitative feature is that the scenarios pick up the asymetry in the prediction interval seen in the stochastic forecasts that is due to the log-normality of population size predicted by stochastic renewal theory (Tuljapurkar, 1992).

3.2 Uncertainty in Age Structure Measures

To produce uncertainty estimates on age-structural measures from scenarios, we combine the uncertainty of different age groups in a probabilistically sensitive manner, taking into account the general magnitude of the correlations between age segments observed in Table 1.

Demographers define dependency ratios as the ratio of dependent population age groups to the working age population. Following convention, the working age population W consists of those aged

Figure 4. *Stochastic forecast of total fertility rate 95% forecast interval for each year and 95% forecast interval for cumulative average.*



The traditional scenario approach uses within-scenario age group sizes. The adjusted approach takes account of the possibility that all age group sizes may not be "large", or "small", at the same time.

20 through 64, the young population Y of those aged 0 through 19, and the elderly population O of those aged 65 and above. Then, the old-age dependency ratio (ODR) is O/W; the youth dependency ratio (YDR) is Y/W; and the total dependency ratio (TDR) is $(Y+O)/W$, the sum of the youth and old-age dependency ratios.

Our strategy is to estimate the uncertainty in dependency ratios using the standard approximation for the variance of correlated random variables (e.g. Rice (1995)),

$$\sigma^2_{X/Y} \approx \left(\frac{\mu_X}{\mu_Y}\right)^2 \left(c_X^2 + c_Y^2 - 2\rho_{XY} c_X c_Y\right), \tag{1}$$

where μ_X and μ_Y are the means of X and Y, c_X and c_Y are the coefficients of variation (the ratio of the standard deviation to the mean), and ρ_{XY} is the correlation between X and Y.

Using the exact values of the correlation observed between age groups would produce the best estimate of uncertainty. In general, however, these will not be known, and so we assume correlations of 1 between the young and working age populations and 0 between the elderly and non-elderly populations. We estimate the coefficients of variation, assuming normality, as one-fourth the ratio of the width of the prediction interval to the central forecast.

Figure 4 shows the dependency ratios of the stochastic approach and those produced by the modified scenarios. The figure also shows the dependency ratio ranges produced by taking the range of dependency ratios directly from the scenario output.

In the case of the old-age dependency ratio, the modified scenario approach produces an uncertainty interval that tracks the stochastic intervals nearly exactly for the first 40 to 50 years of the forecast, and does so much better than the direct use of scenarios. After this, the width of the modified scenario intervals are comparable to the stochastic intervals but the use of the normal approximation does not capture the asymmetry of the stochastic forecast intervals. Still, allowing fertility and mortality to vary independently produces much more accurate uncertainty intervals than does the traditional use of scenarios.

In the case of the youth dependency ratio, our assumption of perfect correlation between the young and working age populations is equivalent to the traditional use of scenarios. Both scenario approaches produce somewhat narrower uncertainty bounds than they should. For example, using the observed correlation between the young and the working age segments given in Table 1 would increase the uncertainty by a factor of almost a third in the 70th year of the forecast. Still, our general finding is that the bounds on the youth dependency ratio, even assuming perfect correlation, are of the same general magnitude as those produced by the stochastic forecast.

Finally, we turn to the total dependency ratio, which requires an additional assumption to be made about the covariance between the youth and elderly dependency ratios, since $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X, Y)$. The fully stochastic results (not shown) reveal that in these forecasts the old-age and youth dependency ratios are *negatively* correlated, although it is not clear that this is a general property across populations. Our estimates of uncertainty in the old-age and youth dependency ratios are both on the low side, so assuming no covariance between the two ratios offsets this bias. The result is a remarkably close match in the uncertainty intervals from the stochastic and the modified scenario forecasts. The improvement relative to the traditional use of scenarios is dramatic.

4 Adjusting Dependency Ratios from Traditional Scenarios: A Worked Example

In applications of the method proposed here, one would not generally have estimated versions of the fertility and mortality time series—if one did, one could just continue with the fully stochastic forecast. The more likely real-world situation is that one has the output from scenarios and wants to present these in a probabilistically sensible manner. To do this, one would take the output for each age group from the scenarios, use the bounds generated for each age group and total population directly from the scenarios, but recomputing the uncertainty in the total dependency ratios using the methods proposed here.

As an example, Table 2 presents the projected populations in year 50 of the forecast (year 2040) based on the scenarios used in this paper. The coefficients of variation estimated as (Large–Small)/(4 $\times$ Medium), based on assuming normality with the Medium scenario as mean. We then apply equation 1 assuming $\rho = 0$ for the ODR and $\rho = 1$ for the YDR. The TDR is calculated by assuming independence between the ODR and YDR, making the variances additive.

The table shows the original and adjusted dependency ratios along with the ranges calculated using equation 1. We see that the YDR is unaffected by the adjustment, since the scenarios implicitly assume correlations of 1 between the young and working age groups. The uncertainty range in the old-age dependency ratio, on the other hand, increases substantially, some 5-fold, due to our adjustment. This is because the original scenarios always combine large working populations with large elderly populations and small working populations with small working populations, whereas in fact independence—or something very close to it—is warranted. The adjusted uncertainty in the total dependency ratio is also larger than the uncertainty implied by the original scenarios. Here the adjustment increases the estimate of uncertainty by about a factor of 1/4. The adjusted TDR differs

Table 2

Illustrative example of calculation of adjusted dependency ratios.

(a) Age-group results of year 50 of scenario forecasts

	Millions of females			
Age group	Small	Medium	Large	Coef. of Var.
0–19	36.7	18.4	62.5	0.30
20–64	78.9	64.2	94.8	0.10
65+	32.2	27.9	36.5	0.07

(b) Original and adjusted dependency ratios

	ODR		YDR		TDR	
	original	adjusted	original	adjusted	original	adjusted
Low	0.39	0.31	0.29	0.29	0.72	0.66
Medium	0.41	0.41	0.47	0.47	0.87	0.87
High	0.43	0.51	0.66	0.66	1.04	1.08

Coefficients of variation estimated using normal approximation as one-fourth the difference between Large and Small divided by Medium. “Original” dependency ratio uncertainty is directly from scenario outputs. “Adjusted” uncertainty range between “Low” and “High” is the 95% forecast interval based on equation 1 and normal approximation. Adjusted TDR assumes independence of YDR and ODR.

from the original TDR by larger amounts later in the forecast (See Figure 4).

This example illustrates how to use the scenario results, conditional on having scenarios that themselves have some probabilistic interpretation. In real applications, one would need to specify the scenarios. Alternatives to stochastic time series could involve the use of judgement and expert opinion (Lutz *et al.*, 2001) and/or *ex post* evaluation of past scenarios. These approaches are discussed further below.

5 Discussion

In this section, we first discuss the need for accurate estimates of uncertainty and then consider ways to implement the modified scenario approach using stochastic time series methods, expert opinion, and existing scenario-based forecasts.

5.1 How Accurate do Uncertainty Measures Need to Be?

The modified scenario approach produces uncertainty intervals that are broadly similar to those produced by the fully stochastic forecasts. But differences emerge, particularly for forecast horizons beyond about 60 years when fertility uncertainty begins to influence the number of elderly.

How accurate do estimates of uncertainty need to be? It is sometimes argued that estimates of uncertainty from the fully stochastic forecasts are misleadingly precise. Exact intervals are produced but these are conditional on the correct specification of the time series model and more importantly on the assumption that past variability is a good estimate of future uncertainty. In such a situation, approximations of stochastic forecasts such as those produced here may arguably give a somewhat more appropriate, even if less precise, depiction of the uncertainty in uncertainty estimates.

At a practical level, the degree of uncertainty at longer time horizons is often of less interest to contemporary planners because advance warning will come if unexpected changes take place. For example, a baby boom in the decade of the 2020s would surprise most contemporary observers. But because it will take some years for these newborns to enter school, even more years for them to enter the workforce, and many decades for them to enter the retirement system, there will be advance notice, which new forecasts can then take into account. Even when a long planning horizon is legally mandated, as is the 75 years for actuarial reports of the U.S. social security system, policy debates usually revolve around changes expected in the next several decades. A final factor giving short horizons greater importance is that policy makers recognize that the distant future is more uncertain than the near future and are thus less likely to lock in inflexible policies with extremely long time horizons. For all of these reasons it is important to emphasize the success of the modified scenarios during the first five decades of the forecast.

5.2 Calibrating Scenarios

Here, I have calibrated scenarios to the prediction intervals generated by stochastic time-series models. When time-series models for vital rates are available, it will usually be worth continuing with fully stochastic forecasts of the population, with calibrated scenarios best used as a tool for checking stochastic forecasts.

The modified scenario approach is most useful when time series estimates are either unavailable or, for whatever reason, unbelievable. In these instances, one would want to calibrate the scenarios to some other standard. Two approaches seem reasonable. The first is a minor modification of existing scenario methods whereby experts either within or outside of forecasting agencies specify high and low scenarios. In this case, the results presented here suggest that there be an explicit statement by those designing fertility scenarios that the scenarios apply to long-term averages, not to the uncertainty in any given year. The stability of the uncertainty interval of the averages over a broad range of cumulative averages from 20 to 80 years reduces the need for experts to worry about the precise form of the average they are considering, as long as it is made clear that it is at least several decades.

A second approach would use the effective uncertainty range covered by past scenarios and apply it to the current production of scenarios. Stoto has found that forecasts in the industrial world tend to cover about a two-thirds prediction interval in population size in any given year (Stoto, 1983). The present analysis suggests doing an *ex post* analysis exercize but with a focus on the cumulative average of fertility. The quantile range covered by past scenarios could then be used to give a probabilistic interpretation to the results from new scenario-based projections, extrapolating these *ex post* errors into the future.

Scenario forecasts could also be made without specifying the probability interval covered by “high” and “low” but by using the methods of combining age segments and calculating the uncertainty in dependency ratios. This would assure greater consistency in the uncertainty range of population age group sizes, total population, and dependency ratios, so that all would cover approximately the same uncertainty interval, even if it remained undefined.

6 Conclusion

We have found that scenarios, when used with some care, can come close to reproducing the uncertainty estimates produced by fully stochastic population forecasts. The key elements are the specification of a probability interval covered by the scenarios, the calibration of fertility scenarios to variability in long-term average fertility, and the probabilistic combination of uncertainty from distinct age segments.

Several caveats are in order. The first is that the uncertainty estimates from our use of scenarios came closest to the estimates from stochastic scenarios for the first 60 years or so. For long term forecasting, full incorporation of stochastic renewal becomes more important. Second, our results apply to populations in which future fertility and mortality can be assumed to be independent and where there is little mortality at younger ages. The probabilistically sensitive use of scenarios in populations in which mortality effects the size of the young and working age populations needs further research. Third, the methods used here apply only to broad age segments. They are inappropriate for assessing the relationships between the joint uncertainty of functions of narrower age groups. For uncertainty estimates for detailed age-distributions, the fully stochastic approach is needed. Finally, further research is needed to assess uncertainty in time-integrated functions of population structure, such as the viability of pay-as-you-go pension systems. Unlike the scenarios presented here, the fully stochastic approach can be integrated over time without concerns about probabilistic consistency.

Overall, our results suggest some interesting future applications of probabilistically sensible scenario forecasts. It may be possible to translate existing scenario forecasts into approximate probabilistic forecasts fairly easily. It also suggests a method for translating expert-based forecasts into probability intervals. Finally, the approach offers a useful first step for those agencies considering the production of probabilistic forecasts using the fully stochastic time-series approach.

Acknowledgement

I thank Gyanendra Badgaiyan for research assistance and am grateful for comments on an earlier version of this paper to Juha Alho, Ronald Lee, Tim Miller, Guy Stecklov, Shripad Tuljapurkar, Warren Sanderson, and participants of the Stochastic Forecasting Workshop held in December 2002 at the Vienna Institute for Demography.

References

Alho, J.M. & Spencer, B.D. (1991). A population forecast as a data base: implementing the stochastic propagation of error. *Journal of Official Statistics*, **7**, 295–310.

Alho, J.M. (1992). The magnitude of error due to different vital processes in population forecasts. *International Journal of Forecasting*, **8**, 301–314.

Cohen, J.E. (1986). Population forecasts and confidence intervals for Sweden: a comparison of model-based and empirical approaches. *Demography*, **23**, 105–126.

Congressional Budget Office (2001). Uncertainty in social security's long-term finances: a stochastic analysis. CBO Paper, December 2001, Congress of the United States.

Keilman, N., Pham, D.Q. & Hetland, A. (2002). Why population forecasts should be probabilistic—illustrated by the case of Norway. *Demographic Research*, **6**, 410–453, Article 15.

Keyfitz, N. (1972). On future population. *J. Amer. Statist. Assoc.*, **67**(338), 347–363.

Lee, R.D. (1993). Modeling and forecasting the time series of U.S. fertility: age distribution, range, and ultimate level. *International Journal of Forecasting*, **9**, 187–202.

Lee, R.D. (1999). Probabilistic approaches to population forecasting. In *Frontiers of Population Forecasting*, Vol. **24** of *Supplement to Population and Development Review*, Eds. W. Lutz, J. Vaupel and D. Ahlburg, pp. 156–190. New York: Population Council.

Lee, R.D. & Carter, L. (1992a). Modeling and forecasting the time series of U.S. mortality. *J. Amer. Statist. Assoc.*, **87**, 187–202.

Lee, R.D. & Carter, L.R. (1992b). Modeling and forecasting U.S. mortality: rejoinder. *J. Amer. Statist. Assoc.*, **87**(419), 674–675.

Lee, R.D. & Tuljapurkar, S. (1994). Stochastic population forecasts for the United States: beyond high, medium, and low. *J. Amer. Statist. Assoc.*, **89**(428), 1175-1189.

Lee, R.D. & Tuljapurkar, S. (1998). *Stochastic forecasts for social security*. In *Frontiers in the Economics of Aging*, Ed. D. Wise, pp. 393–420. Chicago: University of Chicago Press.

Lutz, W. & Scherbov, S. (1998). An expert-based framework for probabilistic national population projections: the example of Austria. *European Journal of Population*, **14**, 1–17.

Lutz, W., Sanderson, W. & Scherbov, S. (2001). The end of world population growth. *Nature*, **412**, 543–545.

National Research Council (2000). *Beyond Six Billion: Forecasting the World's Population*, Eds. J. Bongaarts and R.A. Bulatao. Washington, DC: National Academy Press.

Rice, J.A. (1995). *Mathematical Statistics and Data Analysis*, second edn. Belmont, California: Duxbury.

Stoto, M. (1983). The accuracy of population projections. *J. Amer. Statist. Assoc.*, **78**, 13–20.

Tuljapurkar, S. (1992). Stochastic population forecasts and their uses. *International Journal of Forecasting*, **8**, 385–391.

Résumé

L'approche traditionnelle à trois scénarios (supérieur-moyen-inférieur) pour quantifier l'incertitude dans les projections de population a été critiquée pour son manque de signification et de cohérence probabilistes. Cet article montre, sous certaines hypothèses, comment on peut utiliser à bon escient des scénarios calibrés pour estimer les intervalles d'incertitude sur la taille et la structure par âge de la population totale, obtenues avec des projections entêrement stochastiques. De nombreuses organisations produisent déjà des scénarios de projections que l'on sait traiter. Aussi les méthodes présentées ici apportent un point de vue intermédiaire intéressant entre l'analyse d'un scénario incohérent au sens probabiliste et les projections entièrement stochastiques.

Appendices

A Stochastic Processes for Vital Rates

All of the estimates in this paper are based on the estimates in Lee & Tuljapurkar (1994) and Lee (1993). The fertility model used is a constrained autoregressive moving average model, ARMA(1,1), given by

$$F_t = cF_{t-1} + F(1-c)u_t + du_{t-1}, \tag{2}$$

with $c = 0.9676$ and $d = 0.47978$, the ultimate mean level equal to $F = 2.1$ and with $u_t \sim \mathcal{N}(0, \sigma_u)$, $\sigma_u = 0.110663$. This index of fertility is translated to age specific fertility rates according to the schedules in Lee (1993).

The mortality model is a random walk with drift, given by

$$K_t = K_t - z + \epsilon_t, \tag{3}$$

with $z = 0.365$ and $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon)$, $\sigma_\epsilon = 0.651$. This mortality factor is translated to age-specific mortality rates, namely

$$\log m_{x,t} = \alpha_x + \beta_x K_t.$$

The schedules α_x and β_x are given in Lee & Carter (1992a).

Both formulations assume perfect autocorrelation between age-groups from year to year. Estimates of uncertainty omit uncertainty in the estimation of parameters, following Lee & Tuljapurkar (1994).

B A Generational Model of Stochastic Renewal

To represent the stochastic forecast, let X_t be the net reproductive rate of generation t such that

$$G_t = (X_t \cdot X_{t-1} \cdot \ldots \cdot X_1)G_0.$$

The X_t are serially correlated random variables such that

$$X_t = M + \epsilon_t,$$

where M is the expected net reproduction rate and ϵ is a random variable with zero mean, variance σ_t^2 and $\mathrm{cov}(\epsilon_t, \epsilon_{t-1}) \geq 0$.

As an alternative, generation size in the scenario based forecast can be defined as

$$G_t^s = Y^t G_0,$$

where now Y is a single random variable that when multiplied by itself t times will give the change in G_0 that gives G_t.

Letting $Z_t^s = Y^t$ and $Z_t = \prod^t X_i$, we wish to find $E(Y)$ and $\mathrm{var}(Y)$ such that $E(Z_t) = E(Z_t^s)$ and $\mathrm{var}(Z_t) = \mathrm{var}(Z_t^s)$. Using the delta method to first order for the expectation (Rice, 1995), $E(Z_t) \approx E(X_1) \ldots E(X_t) = M^t$ and $E(Z_t^s) \approx E(Y)^t$. Thus $E(Y) \approx M$. The same method, after some algebra, gives a first order approximation of the variance of the generational stochastic model of

$$\mathrm{var}(Z_t) \approx M^{2(t-1)}\mathrm{var}(X_1 + \ldots + X_t).$$

Substituting $Y = X_i$ for all i, the delta approximation of the generational scenario model gives

$$\mathrm{var}(Z_t^s) \approx M^{2(t-1)}\mathrm{var}(tY).$$

We can now find the variance of Y by setting $\mathrm{var}(Z_t) = \mathrm{var}(Z_t^s)$. This gives

$$\mathrm{var}(Y) = \mathrm{var}\left(\frac{X_1 + \ldots + X_t}{t}\right),$$

which shows that calibrating Y to the variance of the cumulative average of X will equalize the variances of generation size.

The generational model differs from that used in applications in several respects. First, in applications the projections are age-structured, with age groups of 1 or 5 years rather than generational. Second, the scenario calibration ends up giving the scenarios Y a different variance for each year of the forecast, although as we saw this stabilizes after about year 20 because of autocorrelation in the process for annual fertility fluctuations. Finally, the application considers both fertility and mortality independently, rather than net fertility, although the distinction in low mortality populations is not substantial.

The simulation results shown in the body of the paper show that these differences do not, in practice, change the approximate equivalence of the two approaches.

[*Received December 2002, accepted November 2003*]