

## Random Scenario Forecasts

## Versus Stochastic Forecasts

Shripad Tuljapurkar, Ronald D. Lee, and Qi Li



Project #: UM01-03

# **“Random Scenario Forecasts Versus Stochastic Forecasts”**

Shripad Tuljapurkar  
Stanford University

Ronald D Lee  
University of California

Qi Li  
Stanford University

January 2004

Michigan Retirement Research Center  
University of Michigan  
P.O. Box 1248  
Ann Arbor, MI 48104

## **Acknowledgements**

This work was supported by a grant from the Social Security Administration through the Michigan Retirement Research Center (Grant # 10-P-98358-5). The opinions and conclusions are solely those of the authors and should not be considered as representing the opinions or policy of the Social Security Administration or any agency of the Federal Government.

## **Regents of the University of Michigan**

David A. Brandon, Ann Arbor; Laurence B. Deitch, Bingham Farms; Olivia P. Maynard, Goodrich; Rebecca McGowan, Ann Arbor; Andrea Fischer Newman, Ann Arbor; Andrew C. Richner, Grosse Pointe Park; S. Martin Taylor, Gross Pointe Farms; Katherine E. White, Ann Arbor; Mary Sue Coleman, ex officio

# **Random Scenario Forecasts Versus Stochastic Forecasts**

Shripad Tuljapurkar  
Ronald D Lee  
Qi Li

## **Abstract**

Probabilistic population forecasts are useful because they describe uncertainty in a quantitatively useful way. One approach (that we call LT) uses historical data to estimate stochastic models (e.g., a time series model) of vital rates, and then makes forecasts. Another (we call it RS) began as a kind of randomized scenario: we consider its simplest variant, in which expert opinion is used to make probability distributions for terminal vital rates, and smooth trajectories are followed over time. We use analysis and C:\Eudora\attach\demo\_3\_25\_04.pdf examples to show several key differences between these methods: serial correlations in the forecast are much smaller in LT; the variance in LT models of vital rates (especially fertility) is much higher than in RS models that are based on official expert scenarios; trajectories in LT are much more irregular than in RS; probability intervals in LT tend to widen faster over forecast time. Newer versions of RS have been developed that reduce or eliminate some of these differences.

## **Authors' Acknowledgements**

This work was supported by grants from the National Institute of Aging to Tuljapurkar and to Lee, and by the Center for the Economics and Demography of Aging at the University of California at Berkeley, and from the Michigan Retirement Research Center. Li was also supported in part by the Morrison Institute of Population and Resource Studies at Stanford University.

## INTRODUCTION

Population forecasts are widely used by governments, and agencies such as the United Nations, as an essential part of policy analysis. A population forecast requires launch data – the size and composition of the initial population, its vital rates – and a procedure for forecasting the population’s vital rates (mortality, fertility, immigration) over the period covered by (also called the span of) the forecast. Forecasters recognize that the demographic basis for forecasting vital rates over time is inadequate, in the sense that any single forecast is likely to be off the mark. In consequence, demographers attempt to incorporate uncertainty into forecasts, usually by generating a range of forecasts along with some indication of their relative likelihood. A widely used method to indicate uncertainty is to prepare three forecasts, called High-Medium-Low, in which the Medium is most likely, and the others indicate the range of less likely outcomes. This is what the United Nations and the United States Census Bureau, for example, do in official projections. In practice there may be many more than three underlying forecasts but they are combined to present three outcomes for any quantity of interest – we call these *scenario* forecasts. The shortcomings of such forecasts are discussed by, for example, Lee and Tuljapurkar (2000).

There are three general ways of improving on scenario forecasts. One method, due to Keyfitz (1981), Stoto (1983) and Keilman (1998), starts with a historical analysis of errors in past forecasts. The statistical properties of errors are assumed to be the same in future forecasts as in the past, so that the historical data may be used to generate prediction errors. A second approach is a *stochastic* forecast as described by Lee and Tuljapurkar (1994) – a statistical analysis of historical data is used to estimate dynamic stochastic models (e.g., a time series model) of the variability in vital rates. A stochastic forecast is then made using these models – here we call it the LT forecast – that yields probabilities for different ranges of any outcome of interest. In effect, the LT method relies upon a probability distribution for the trajectories of the process that is generated dynamically. The main point of the LT method is that a forecast should reflect as directly as possible the uncertainty displayed by historical data. The philosophy of the LT method is based on and shared by other dynamic approaches to generating future uncertainty, notably the work of Alho and Spencer (1985), Cohen (1986), and Alho (1990) who uses in addition a form of *post hoc* error analysis to project predictive probability

distributions.

A third approach due to Pflaumer (1988) and Lutz et al. (1996, 1997) uses a *random scenario* method that we refer to here as an RS forecast. In its original versions, an RS forecast uses historical and other analysis, and expert opinion, to assert that vital rates at a particular future target date (or set of dates) will lie in a specified range and have a particular probability distribution. Expert opinion may be solicited formally: see Lutz et al. (1996) for a discussion and example. To generate an RS forecast, one target value of vital rates at the target date is selected from the assumed distribution, and it is assumed that vital rates will follow some specified smooth trajectory from launch value to target value. A forecast is made using vital rates on this trajectory; a new target is selected, a new trajectory of rates is generated, a new forecast is made; and so on until one has a large set of forecasts. These forecasts have a probability distribution that reflects the forecaster's assumptions about target values, their distribution, and the trajectory of vital rates. This probability distribution is now used as a basis for making probabilistic forecasts for quantities of interest.

The original RS and LT approaches are very different in their approach to formulating the dynamics that underlie their respective forecasts. The objective of this paper is to compare the two approaches in terms of the properties of the forecasts that they generate. We use two examples to compare the original RS and LT approaches – first, a stylized model of a population without age structure, and second, a complete projection model for the US population with a launch date of 1997. The stylized model highlights the differences between the dynamics of the stochastic processes in RS and LT, and the complex model shows how the content of the forecasts differs in various substantive ways. Our practical aim is to educate makers and users of forecasts to the key differences that arise between the two types of forecast, especially in terms of probability distributions of outcomes and dynamic correlations between important variables over time.

Our comparison deliberately employs stylized versions of RS and LT in which any differences between the two approaches are likely to be large. In practice, as pointed out by many forecasters, virtually all forecasts may require elements of both approaches. For example, the LT approach exploits long runs of accurate historical data when they are available, whereas the RS approach is a way of filling in gaps in the data or in the

mechanisms of change. Thus it is natural and sometimes essential to consider intermediate modelling methods. One notable effort in this direction is the work of Lutz et al. (2001) who extended the original RS method by assuming that the vital rate trajectories have an additional random component. Their new method is a substantial advance over the original RS, but we will not discuss it here. Our goal is not to make a comprehensive comparison between actual forecast outcomes using the latest methods, but only to show how different forecast methodologies can generate persistent differences in the forecasts that they yield. Armed with the understanding that we provide here of how these differences arise in simple cases, the reader will be prepared to explore how the newer methods of Lutz et al. (2001) narrow these differences to good effect.

The next two sections compare a simple stylized RS model and an LT model for two cases: first, a simple scalar population dynamics, and second, the more realistic case of projecting the US population. The final section discusses our findings and conclusions.

## SCALAR FORECASTS

We begin with the simplest setting in which to examine how one makes a forecast. Consider a population which has total size  $N(t)$  at the end of year  $t$  – in this section we ignore age structure and other details. During the subsequent year the population is assumed to grow by a ratio  $R(t + 1)$  to a number

$$N(t + 1) = R(t + 1)N(t). \tag{1}$$

To make a forecast starting with a known  $N(0)$  at year  $t = 0$  until the end of a forecast span of  $T$  years, we need only specify the growth factors  $R(1), R(2), \dots, R(T)$ . How do we specify and compare forecasts in the simplest versions of the RS and LT approaches? We present below simple but representative models for the dynamics as used in RS and LT, and examine their outcomes; analytical results are given but the focus is on graphical comparisons using a representative numerical example.

### A Simple RS Model and Forecast

A simple stylized RS forecast is described by assumptions about the final distribution of the growth rates, and the intermediate trajectories.

1. There is a known initial growth factor  $R(0)$  and initial population  $N(0)$ .
2. The growth factor in the last year,  $R(T)$ , has a most likely value  $R_M$ . With high probability (say 90%) the value of  $R(T)$  will lie in the range from  $R_L < R_M$  to  $R_H > R_M$ . In practice, these values can be based on a survey of expert opinion.
3. The value of  $R(T)$  has either a uniform or a normal random distribution with mean  $R_M$  and a variance  $\sigma_R$  that reflects the range assumed above.
4. Given any particular value of  $R(T)$ , the growth factors for all forecast periods are obtained by interpolation between  $R(0)$  and  $R_1$ ,

$$R(t) = (1 - h(t))R(0) + h(t)R(T), \quad (2)$$

where  $h(0) = 0, 0 \leq h(t) \leq 1$ , and  $h(T) = 1$ . We assume a specific interpolation function, e.g.,  $h(t) = (t/T)$ .

To make the RS forecast we simply select a value of  $R(T)$  according to the assumptions, generate a trajectory of  $N(t)$  between  $t = 1$  and  $T$ , and repeat. Our numerical illustrations are based on 500 such trajectories.

### A Simple LT Model and Forecast

A simple LT model for this setting is described by assumptions about a stochastic process that generates trajectories of growth rates over time.

1. There is a known initial growth factor  $R(0)$  and initial population  $N(0)$ .
2. The growth factor  $R(t)$  is a serially correlated random process with an expected mean of  $R_M$  and follows the equation

$$R(t) - R_M = \rho(R(t-1) - R_M) + e(t), \quad (3)$$

where  $e(t)$  is an iid random process, independent of  $R(t)$ , with zero mean and variance  $\sigma_e$ .

3. We specify  $R_M$ , based on either evidence about trends or expert opinion.

4. We specify  $\rho$  and  $\sigma_e$ . In actual application, we would estimate these by fitting the model equation (3) to historical data – this is a fit conditional on the assumed value  $R_M$ .

To make an LT forecast we generate a stochastic sequence of  $R(t)$  for times  $t = 1$  to  $T$  using equation (3), forecast the corresponding sequence of  $N(t)$ , and repeat. Here too, our numerical illustrations are based on 500 such trajectories.

### A Numerical Example

In both the RS and LT cases we obtain many (here, 500) trajectories and then use these samples of the forecast trajectory to generate probability forecasts. For example, if  $T = 50$ , we can describe our forecast of  $N(25)$  for  $t = 25$  in several ways:

1. Plot the 500 trajectories of  $N(t)$  through  $t = 25$ .
2. Use a histogram (or some better statistical density estimator) to estimate the probability distribution of  $N(25)$ .
3. Display a predictive distribution that shows the estimated quantiles of the forecast distribution of  $N(25)$ .
4. Estimate the mean and standard deviation of  $N(25)$ .

In our numerical example, for both RS and LT we set initial values

$$R(0) = 1.028, \quad N(0) = 6.$$

The target for the RS example at  $T = 50$  is chosen to be symmetric with

$$R_L = 0.01, \quad R_M = 0.02, \quad R_H = 0.03.$$

For a normally distributed target we set the RS final variance to be  $\sigma_R = (R_M - R_L)/1.6448$ . For a uniformly distributed target we set  $\sigma_R = (R_M - R_L)/0.45$ .

In the RS example we pick some  $R(T)$  from the target distribution and then connect  $R(0)$  to this final value by an interpolation as in equation (2). The baseline case we use is the **linear** case with  $h(t) = (t/T)$ . We compare this with a **slower** than linear approach to the target with a *quadratic*

interpolation  $h(t) = (t/T)^2$ . We also compare the baseline with a **faster** than linear approach to the target with a *square-root* interpolation  $h(t) = (t/T)^{1/2}$ .

The steady state variance of the LT growth rates is

$$\sigma_{LT}^2 = \sigma_e^2 / (1 - \rho^2). \quad (4)$$

We use values of  $\rho$  from the set  $\{-0.9, -0.5, 0, 0.5, 0.9\}$ . The last value is what we expect for a highly correlated series, as is typical of human fertility rates.

### Making A Comparison

How should we make a sensible comparison between our simple stylized RS and LT models? In a practical case, such as the US forecast in the next section, the two methods use distinct approaches to arrive at the statistical dynamics of growth rates. As a result, the means and variances of the vital rates over time may well be different and in turn will produce differences in the forecasts – we must then keep track of such differences in making any comparison. But in the present simple case we are free to set the parameters in a way that highlights only the essential differences in the dynamics of simple RS and LT models. To keep the models as comparable as possible, we set the LT mean value of growth rate to be  $R_M$ , the same as the RS value, and also make the asymptotic variances equal by setting  $\sigma_{LT} = \sigma_R$  which means that we set

$$\sigma_e = \sqrt{1 - \rho^2} \sigma_R. \quad (5)$$

Given these parameter settings, any differences we find are due solely to the different ways in which the two approaches model the vital rates.

### Comparing Forecast Properties

#### Trajectories

Figure 1 shows 25 trajectories from an RS simulation using a normally distributed target and linear interpolation; also shown are the trajectories that go to the High, Mean and Low target growth rates. Note the smooth trajectories do not cross, and the High-Mean-Low range contains almost all the results.

Figure 2 shows 25 trajectories from an LT simulation with  $\rho = +0.9$ . Note that trajectories do cross, there is a much wider range of population sizes and a “looser” distribution. Instead of a High-Mean-Low we show three trajectories that span the outcomes at the end of the projection, i.e., the trajectories whose terminal values are the median and the upper and lower 95% at the end of the forecast.

### Probability Distributions of Forecast Population

A general feature of RS is that the shape of the target distribution of  $R(T)$ , which is set by assumption, largely determines the shape of the distribution of  $N(T)$ . To see why, note that in RS

$$N(T) = [(1 - h(1))R(0) + h(1)R(T)] [(1 - h(2))R(0) + h(2)R(T)] \dots (6) \\ [(1 - h(1))R(0) + h(T - 1)R(T)] N(0).$$

Hence  $\log N(T)$  has the same distribution as

$$\sum_i \log([(1 - h(i))R(0) + h(i)R(T)]).$$

The factors in this sum are perfectly correlated with each other and the distribution of each will be very similar to that of  $R(T)$  – to obtain each factor we multiply  $R(T)$  by a number and add a number. Therefore the distribution of  $N(T)$  must be similar to the distribution of  $R(T)$ .

This distributional feature of RS is illustrated by Figure 3, in which the top panel is a histogram of the forecasted values of  $N(50)$ , the lower panel is a histogram of  $\log N(50)$ . This figure was generated using a uniformly distributed target, i.e., a uniform distribution for  $R(50)$ , and we see that the result is an approximately uniformly distribution of  $N(50)$ . Note that the logarithm of  $N$  is also approximately uniformly distributed. We note also that if we use instead a normally distributed RS target, we obtain an approximately normal distribution (not shown here) for both  $N(50)$  and its logarithm.

In contrast to RS, a general feature of LT is that the distribution of  $N(T)$  for large  $T$  (say  $T > 10$  or so), is always approximately lognormal. To see why note that in LT we have

$$\log N(T) - \log N(0) = \sum_i \log R(i), \quad (7)$$

where the  $R(i)$  are a serially correlated sequence. But the elements of the sequence are **not** perfectly correlated and in fact the correlation between  $R(t)$  and  $R(t + s)$  goes to zero at the rate  $\rho^s$  as the difference  $s > 0$  increases. This implies that the sum on the right side of the above equation must obey the central limit theorem and be asymptotically normally distributed. In other words,  $\log N(t)$  in LT will be roughly normally distributed and  $N(t)$  will be roughly lognormally distributed. We do not illustrate this feature here.

### **Growth Rate of Variance**

The growth rate of forecast variance in RS depends strongly on the trajectory along which vital rates converge to their target value (e.g., faster or slower than linear) and only weakly on the shape of the target distribution. This is illustrated in Figure 4. To focus on the growth rate of the variance in  $N(t)$  we have scaled out the average growth rate, which in RS is the growth rate along the trajectory towards  $R_M$ . As we might expect, variance grows most rapidly when growth rates converge slowly to the target – in this case different RS trajectories start at the same growth rate but diverge rapidly. This also means that each trajectory is pretty close to its target rate after only a few years. Note that the variances in Figure 4 grow quadratically with time. This too is a feature of the RS projection, as may be seen from equation (7) – the variance of that sum of  $t$  perfectly correlated terms will grow as  $t^2$ .

In contrast consider equation (7) for LT – the variance of a sum of terms whose correlation dies out as the terms become more distant grows linearly as  $t$ . The contrast between forecast variances in RS and LT is shown in Figure 5. The three LT curves show how strongly serial correlation  $\rho$  affects the LT variance, but all three LT curves rise linearly with  $t$  at long times. The RS curve shows its characteristic quadratic rise. Note also that the small  $t$  variances change very differently – the LT variances rises quite rapidly with time at first, then drops to a linear change. The RS variance always grows quadratically. Thus early into the forecast the LT method will usually result in substantially greater forecast variance than will RS.

## AGE-STRUCTURED FORECASTS FOR THE US

We must first consider the appropriate setting in which to compare forecasts. In the scalar model of the previous section, we were free to set the parameters to minimize the difference between forecast methods. In contrast, here we are describing an actual country and demographic history for which both the simple RS method and the LT method must be based on separate analyses that result in dynamic models of mortality and fertility change. For example, the LT model below fits a time series model for fertility but constrains the long-run average fertility level. We constrain this average to equal the ultimate medium fertility scenario of the US Social Security Administration (a similar value is used by the US Census Bureau). The LT time series model generates fertilities that vary substantially around the constrained average. These official agencies also provide a range of high (optimistic) and low (pessimistic) values for the ultimate fertility, and in our simple RS model we assume that they would cover the ultimate probability range (95 percent in case of a normally distributed target). The trajectories of the simple RS are then bounded by that range at all times, and one can only change the variance by changing the target range – this is certainly possible but for our example we rely on the ranges used by official agencies. Thus here we cannot and do not force an equality in the long-run variances of the LT and RS methods. On the other hand, we do impose long-run similarities in the behavior of average fertility in the two approaches, because both require the assumption of a target. We discuss these issues further in the next section.

### LT Forecast

The LT forecast for the US is the one used by Anderson, Tuljapurkar and Lee (2001) with launch data in 1996. Mortality follows a Lee-Carter model estimated using data from 1933 to 1996, fertility follows a time series model using data for the same period (for details consult the reference cited or Lee and Tuljapurkar 1994), and the model trajectories for fertility and mortality are not correlated with each other. We note that fertility displays very high positive serial correlation and that the long-run variance in the total fertility rate projected by LT is high. Immigration is deterministic and follows the US Social Security Administration (1996) medium scenario, hereafter called the SSA medium scenario. The forecast span is 1997 to 2096.

## RS Forecast

We first set assumptions for a stylized version of RS, based loosely on the conventional (not random) scenario methods used by the SSA. Our simple RS model is based on the early work of Pflaumer (1983) and Lutz et al. (1996) and is not representative of the newer RS methods in Lutz et al. (2001). The TFR in the target year 2096 is taken to have High, Medium, Low values of 2.2, 1.9, 1.6, respectively. The Medium to Low difference is used to set the variance of a normally distributed target TFR. The age distribution of fertility is taken to be fixed over time at the value in the launch year. We choose a target in 2096 at random from this distribution and use linear interpolation to generate TFR values for all the launch years: we do this for every choice of target TFR.

Mortality targets are set in terms of targets for the value in 2096 of the expected lifespan at birth, denoted by  $e_0$ . We start with SSA scenarios for 2075 and extend them by interpolation to get scenarios for 2096:

$e_0$	Male 2075	Female 2075	Combined 2075	Combined 2096
High	83.1	88.2	85.7	88.3
Medium	79.3	84.2	81.8	83.3
Low	76.4	81.2	78.8	79.5

We use the difference between Medium and Low to set the variance of a normally distributed target  $e_0$  for each sex. To make population projections we need the full age-specific death rates – we project these in four steps. One, choose a random  $e_0(T)$  target for each sex. Two, assume that age-specific central death rates  $m(x, t)$  at  $t = 0$  and at  $t = T$  are proportional,  $m(x, t) = \lambda m(x, 0)$ . Three, since we know  $m(x, 0)$  we can use  $e_0(T)$  to compute a proportionality factor  $\lambda(T)$  such that the final death rates yield the desired expectation of life. Four, use linear interpolation between  $m(x, 0)$  and  $m(x, T)$  to find the death rates in any forecast year: do this for each sex along each trajectory.

## Forecasts of Population

Any forecast of an age and sex structured population over a long span contains a huge amount of information. We choose only a few components of the forecast for our comparison, aiming to identify the key differences

that obtain between LT and simple RS methods. Where relevant we indicate the substantive importance of the variables we have chosen to discuss.

We begin with the dramatic difference between LT and RS forecasts in terms of the width of their prediction intervals for total population over the entire forecast span, shown by the 95% prediction intervals in Figure 6. Note that the LT forecast is substantially more uncertain even very early into the forecast. We note that in this example the RS forecast of population size is roughly normal because the underlying vital rates have normally distributed targets.

Next we consider (Figure 7) projections of one component of population, the age class 0 to 1 year – births. Births matter to numerous policy issues, ranging from expenditures on maternal care to education. The upper left panel compares the mean of the LT and RS forecasts. Note that the RS forecast shows a “classical” demographic convergence with damped waves that die out by 2050. The mean LT forecast shows substantial variation over time reflecting predicted future baby booms. Although the scale of the upper left panel makes the predicted means of the two methods look fairly close, there are differences of the order of several hundred thousand births over the entire forecast – see the logarithmic plot in the lower left panel. The upper right panel of Figure 7 shows the enormous difference in variance that we have mentioned before. The lower panels of Figure 7 make the same comparison between the logarithm of population in the age group 0 to 1 year. Because the distributions of RS and LT forecasts are very different (RS will be roughly normal, LT will be roughly lognormal) the logarithmic means show a larger difference than the arithmetic means.

Figure 8 shows that the variance difference propagates into other important quantities – here the old age dependency ratio(ODR), defined as the ratio of the numbers of people over 65 years of age to the number of people between ages 20 and 64. The ODR is key to the management of pension policy, in which context it is a plausible proxy for the tax rates needed to support the elderly. The first notable point in Figure 8 is the close agreement between the RS mean forecast and the LT median – this means that even though we did not force assumptions in the two methods to agree, the average values of their forecasts are often close, as is the case here for ODR and to a lesser extent for the number of births. The second point to note is the enormous difference between the uncertainty of the

forecasts – a difference that has implications for any policy analysis. Finally, it is interesting that the RS scenario shows very tight bounds, similar to what the US Census Bureau was reporting about a decade ago.

The differences in forecast variance we have found above for total population, births, and ODR are found in every other projected variable that we examined, for example, the young dependency ratio (YDR). But the similarity between mean forecasts also persists across these variables.

### **Temporal Correlation in Forecasts**

As pointed out earlier (Lee 1999), variables in a scenario or simple RS forecast are typically highly correlated over time along trajectories, and this tends to lock the forecasted variables into highly correlated trajectories. We illustrate this observation with two examples.

Figure 9 shows the correlation between the values of  $e_0$  (sexes combined) and TFR for both forecast methods in 2050. Note that the RS values fall in a tight grouping – indeed they are, as they should be, a bivariate normal distribution with larger variance along the fertility axis. The LT values are substantially more widely distributed: although TFR will be normally distributed in LT the distribution of  $e_0$  is not normal.

Finally, Figure 10 shows estimates of the serial autocorrelation along trajectories of the ODR, estimated using standard time series estimators. Notice that even after 90 years the RS forecasts have a correlation of nearly 0.8 with their starting values. The serial correlation of the LT forecasts falls much more rapidly, declining to about 0.3 in about 25 years.

## **DISCUSSION AND CONCLUSIONS**

We used two forecasting contexts, one simple and stylized, one complex, to compare the probabilistic forecasts that result from a simple random scenario (RS) approach and a stochastic modelling (LT) approach. Our goal has been to show what dynamic and substantive differences result from the differences in the formulation, estimation and dynamics of the models that generate the forecasts.

In our stylized scalar model, we enforced equality of both the long-run averages and long-run variances of the vital rate (the population growth rate). This means that the differences we observed are driven entirely by the different ways in which uncertainty propagates in the RS and LT

formulations of forecast dynamics. We found some key differences. First, the simple RS method yields predictive distributions whose shape is determined by the forecaster’s assumptions about the shape of the target distribution of the random scenarios. In contrast, LT usually generates lognormal distributions of population and its components. Second, the rate of growth of forecast variance in RS is generally quadratic with time at long forecast spans, whereas LT yields a variance that increases linearly with time. Third, at short times into the forecast, the forecast variance initially grows faster in LT than RS.

In our more complex example of the United States population, we construct a simple RS model and contrast it with an LT model that we previously developed. The long-run average vital rates in the two models are similar, but the variances in the long-run are different as dictated by the different models in the two approaches and their estimation. We noted one useful similarity and some important differences. The useful similarity is that the averages of many forecast quantities (e.g., old-age dependency ratio, total population) are quite close in the two approaches. Presumably this a reflection of the similarity in long-run average values of vital rates. One major difference we found was in the level of uncertainty as measured by the width of the predictive distributions of all quantities – the LT method yields greater uncertainty than RS over most of the forecast span with the difference increasing rapidly as we go more than say 25 years into the forecast. The driving force here appears to be the much larger variance of fertility in the LT model. The second major difference is driven by the high positive correlation along forecast trajectories in the RS model: this correlation persists over many decades, in contrast to LT in which the correlations damp out within a generation length.

We note that the Lutz et al. (1996) projections assume much larger variance in TFR than do we in our simple RS method; indeed, they have in 2010 for North America a range of 1.4 to 2.3 to cover 90 percent of a normal distribution. As noted, we relied on officially defined scenarios to obtain a range from an “optimistic” high to a “pessimistic” low. On the other hand, Lutz et al. (1996) use scenario arguments to arrive at the very different view that target fertility values lie in a much wider range. Indeed, we would argue that the use of a much more variable scenario by Lutz et al. supports a point we make – that simply “randomizing” a non-stochastic high-low scenario will not serve to capture the real variability of vital rates.

The RS methods we use here are a “pure” random scenario method

that has since been extended by Lutz et al. (2001) and there are related developments by others (e.g., Pflaumer 1992). The LT methods that we use here are not purely derived by estimating stochastic models; indeed, as we have said, they require the imposition of long-run constraints on the average of fertility. The direction of new developments is towards hybrid methods that combine elements of the simple RS methods we use here with the LT method and related stochastic models. Our goal has been to provide a stylized but sharp comparison of the consequences of the two approaches when they are not hybridized. Makers and users of forecasts should be aware of these differences and the degree to which they may influence the content and interpretation of any probabilistic forecast. We believe that hybrid methods are a logical and important direction in the business of making better and more useful forecasts.

## LITERATURE CITED

- Alho, J.M. 1990. Stochastic methods in population forecasting. *International Journal of Forecasting* 6: 521-530.
- Alho, J.M., and B.D. Spencer. 1985. Uncertain population forecasting, *Journal of the American Statistical Association* 80(390):306-314.
- Anderson, M., S. Tuljapurkar, and Ronald Lee. 1999. Chances are ... Stochastic forecasts of the social security trust fund and attempts to save it. Paper presented at the 1999 Conference on Retirement Research, Center for Retirement Research, Boston College. Available on the web site of the Michigan Retirement Research Center.
- Cohen, J. 1986 Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches. *Demography* 23(1):105-126.
- Keilman, N. 1998. How accurate are the United Nations world population projections? *Population and Development Review* 24(Supplement):15-41,
- Keyfitz, N. 1981. The limits of population forecasting. *Population and Development Review* 7(4): 579-593.
- Lee, R. and S. Tuljapurkar. 1994. Stochastic population forecasts of the U.S.: Beyond High, Medium, Low . *Jour. Amer. Stat. Assoc.* 89: 1175-1189.
- Lee, R. and S. Tuljapurkar. 2000. Population forecasting for fiscal planning: Issues and innovations. IN, Alan Auerbach and Ronald Lee, eds., *Demography and Fiscal Policy*. Cambridge University Press, pp. 7-57.
- Lee, R. 1999. Probabilistic Approaches to Population Forecasting, in Wolfgang Lutz, James Vaupel and Dennis Ahlburg, eds, supplement to v. 24 of *Population Development Review*, 1999 *Rethinking Population Projections*, pp.156-190.
- Lutz, W. W. Sanderson and S. Scherbov. 1996. Probabilistic population projections based on expert opinion. Pages 397-428 in W. Lutz, ed. *The Future Population of the World. What Can We Assume Today?* London: Earthscan, Revised Edition
- Lutz, W. W. Sanderson and S. Scherbov. 1997. Doubling of world population unlikely. *Nature*, Vol. 387, 19 June 1997, pp. 803-805.

Lutz, W. W. Sanderson and S. Scherbov. 2001. The end of world population growth. *Nature* 412:543-545.

Pflaumer, P. 1988. Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting* 4: 135-142.

Pflaumer, P. 1992. Forecasting US population totals with the US Box-Jenkins approach. *International Journal of Forecasting* 8(3): 329-338.

Stoto, M. 1983. The accuracy of population projections. *Journal of the American Statistical Association* 78(381): 13-20.

## FIGURE LEGENDS

FIG. 1. A subset of trajectories (solid lines) of the RS version of a scalar projection, also shown (dashed lines) are trajectories based on the high, medium, low targets.

FIG. 2. A subset of trajectories (solid lines) of the LT version of a scalar projection, also shown (dashed lines) are 95 percent prediction intervals based on 500 simulated trajectories.

FIG. 3. Histograms of 500 forecast values of population in a scalar projection RS model that uses a uniform distribution of the target growth rates in the final year. Upper panel shows the population, lower panel the logarithm of population.

FIG. 4. Forecast variance as a function of forecast span in in a scalar projection RS model. The three pairs of curves obtain for slow, linear, or fast approaches to the target. The choice of target distribution does not matter: solid lines are for uniform target, dashed lines are for a normal target.

FIG. 5. Forecast variances compared for RS and LT in a scalar projection. Variance for LT is shown for three values of serial correlation  $\rho$  – solid line for  $\rho = +0.9$ , dotted line for  $\rho = 0$ , and dash-dot line for  $\rho = -0.5$ . Variance for RS is shown by the dashed line. See text for further discussion.

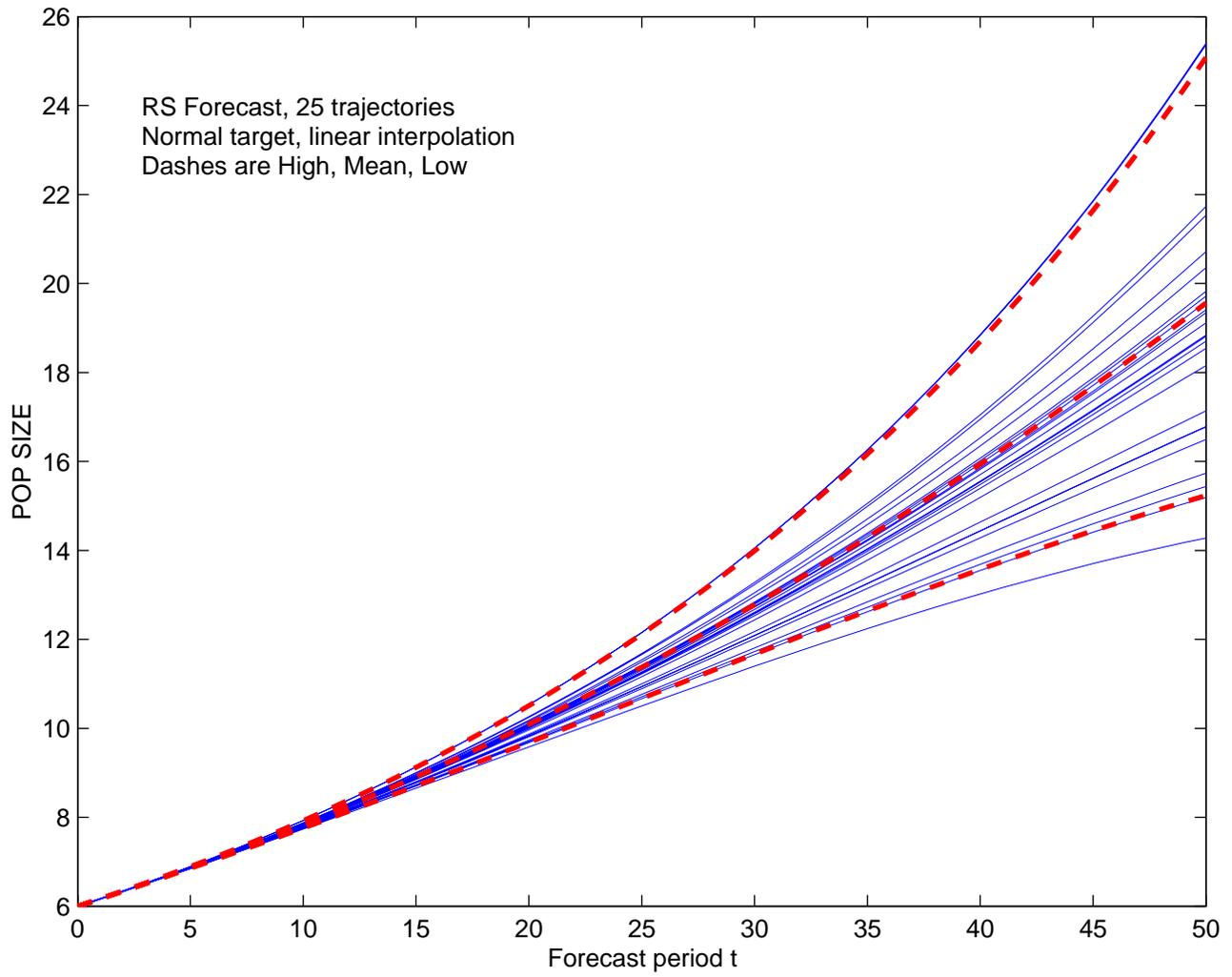
FIG. 6. Percentiles showing the 95 percent predictive distribution of total population in RS (dashed) and LT (solid) projections for the United States, using age and sex specific projection models.

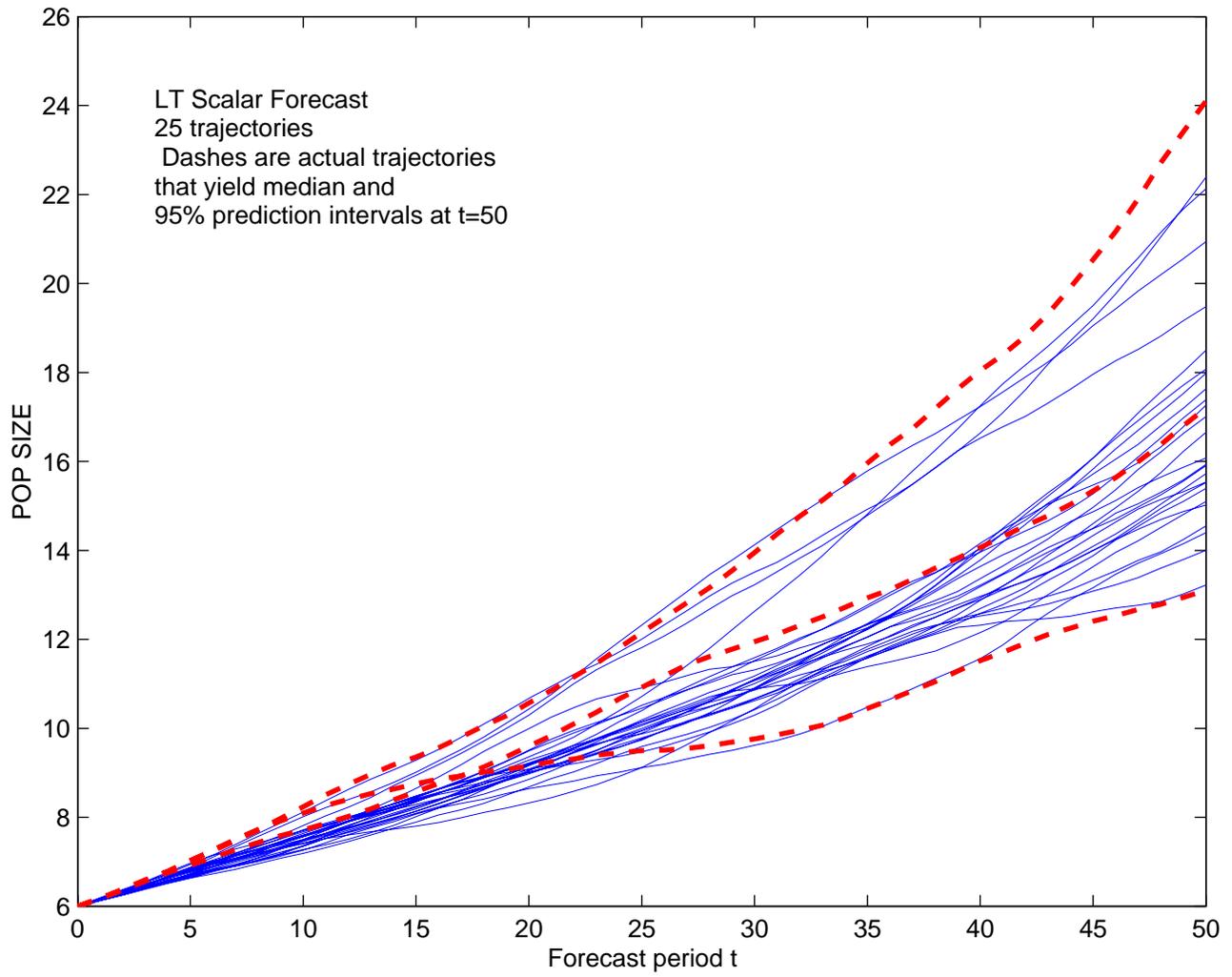
FIG. 7. Averages and variances of births forecast by RS (solid) and LT (dashed) projections for the United States. Upper panels show mean and variance for numbers of births, lower panels for the logarithm of the number of births.

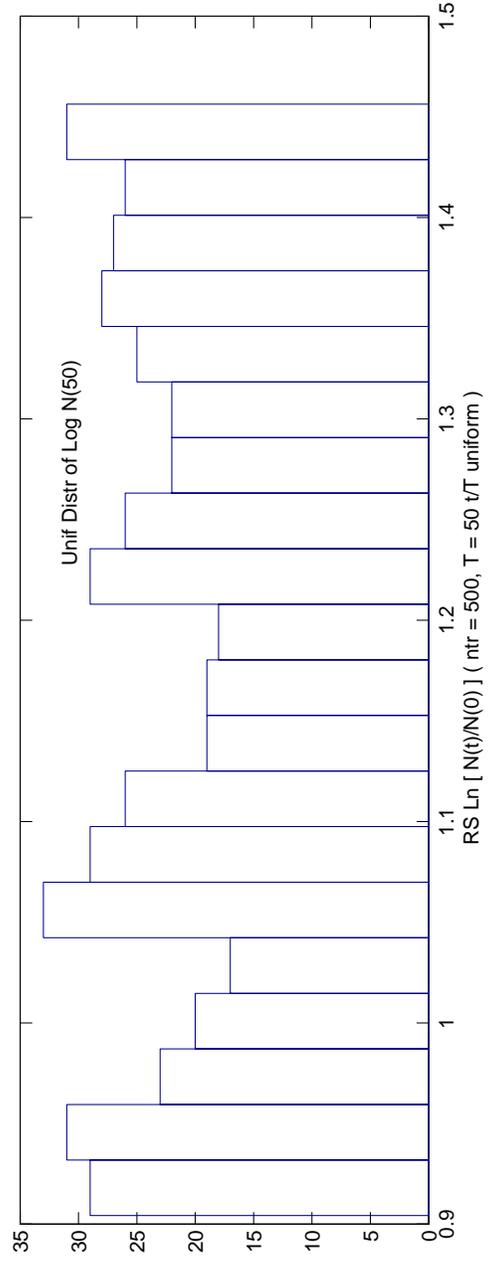
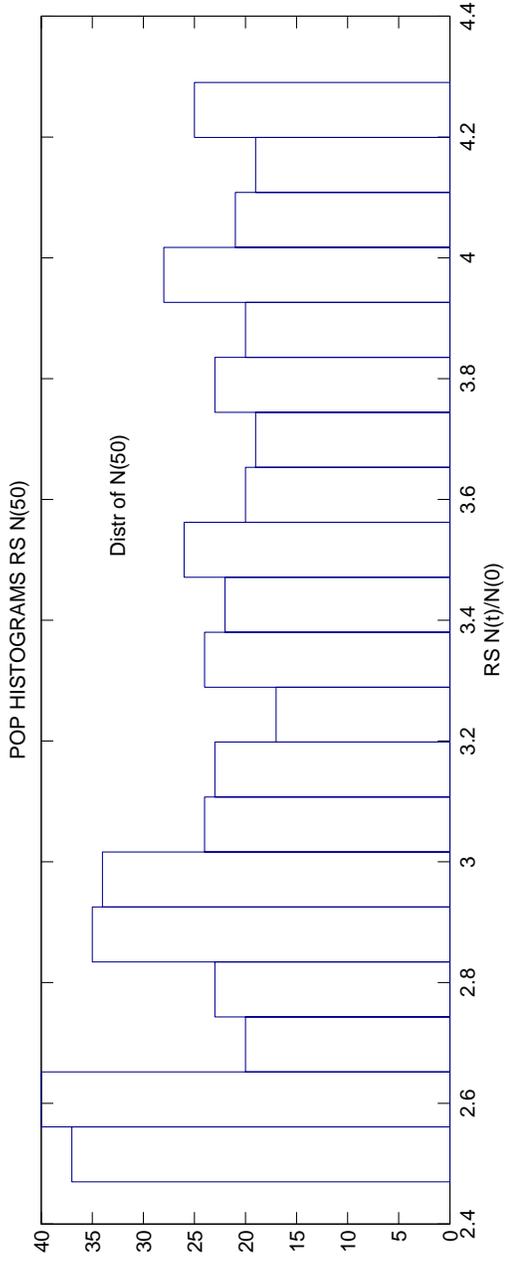
FIG. 8. Percentiles showing the 95 percent predictive distribution of the old age dependency ratio in RS (dashed) and LT (solid) projections for the United States, using age and sex specific projection models.

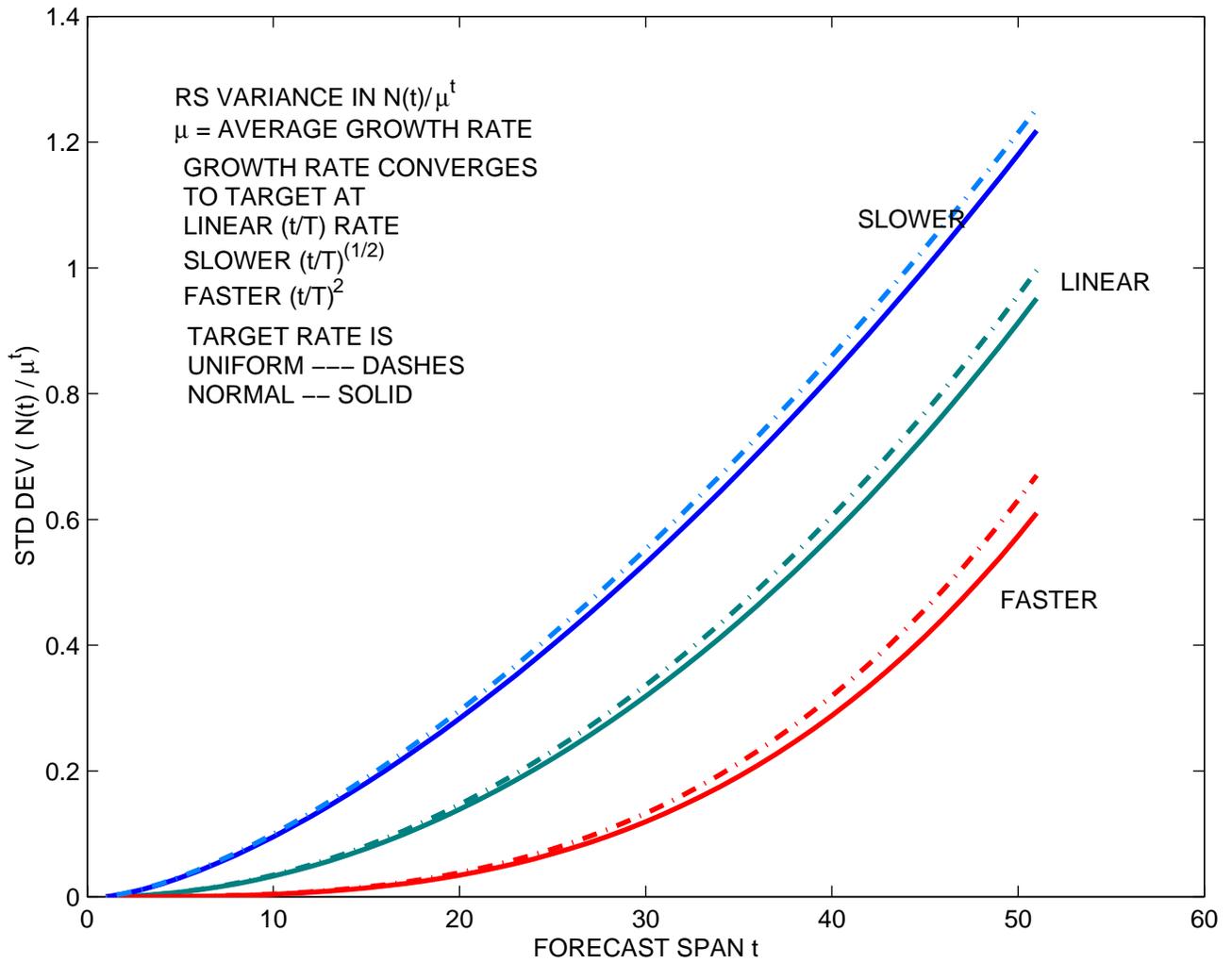
FIG. 9. Scatter plot showing simulated future values in 2050 of life expectancy  $e_0$  and total fertility rate in RS (triangles) and LT (circles) models for the United States.

FIG. 10. Serial autocorrelation of the old-age dependency ratio as forecast by RS (dashes) and LT (solid) models for the United States.

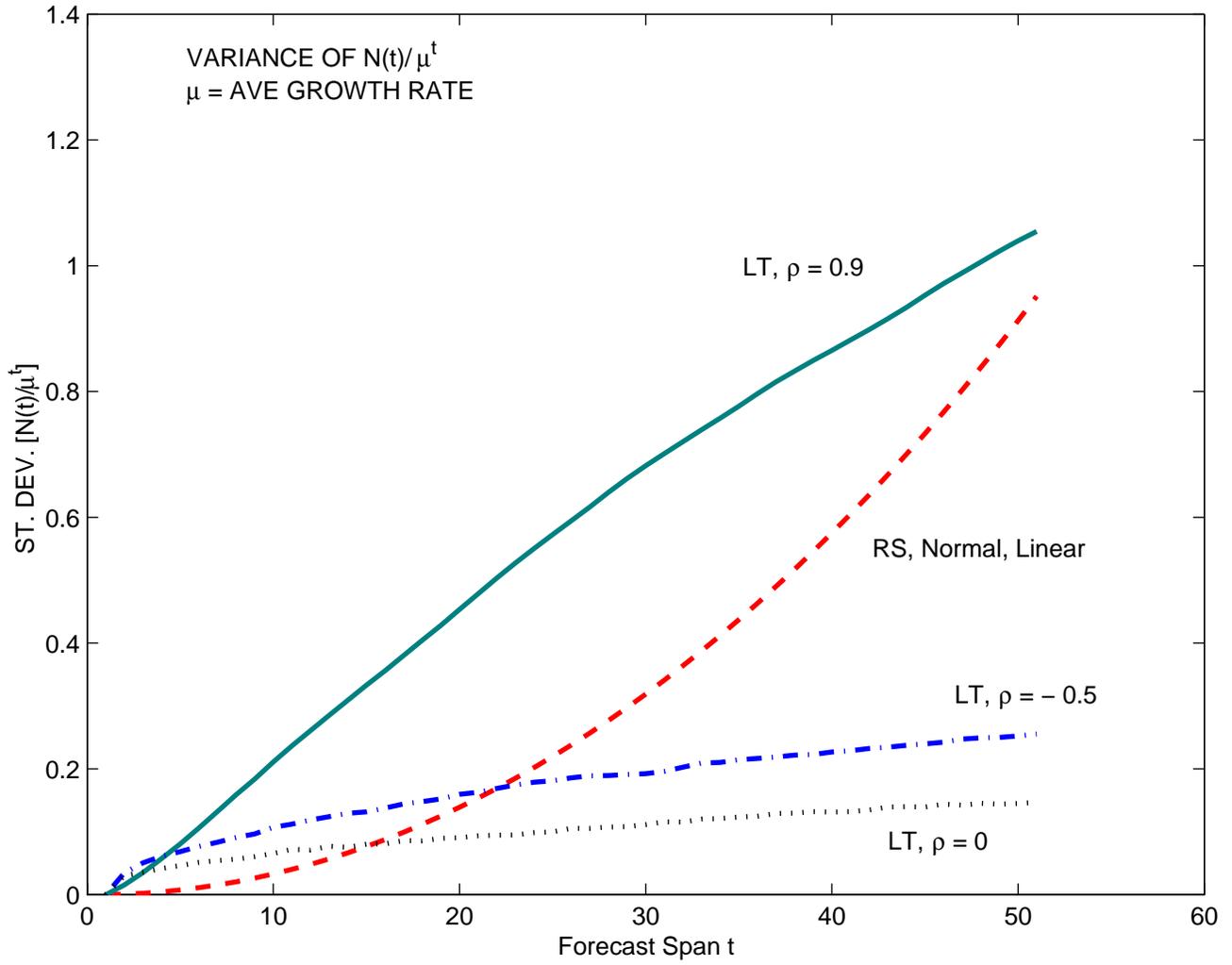


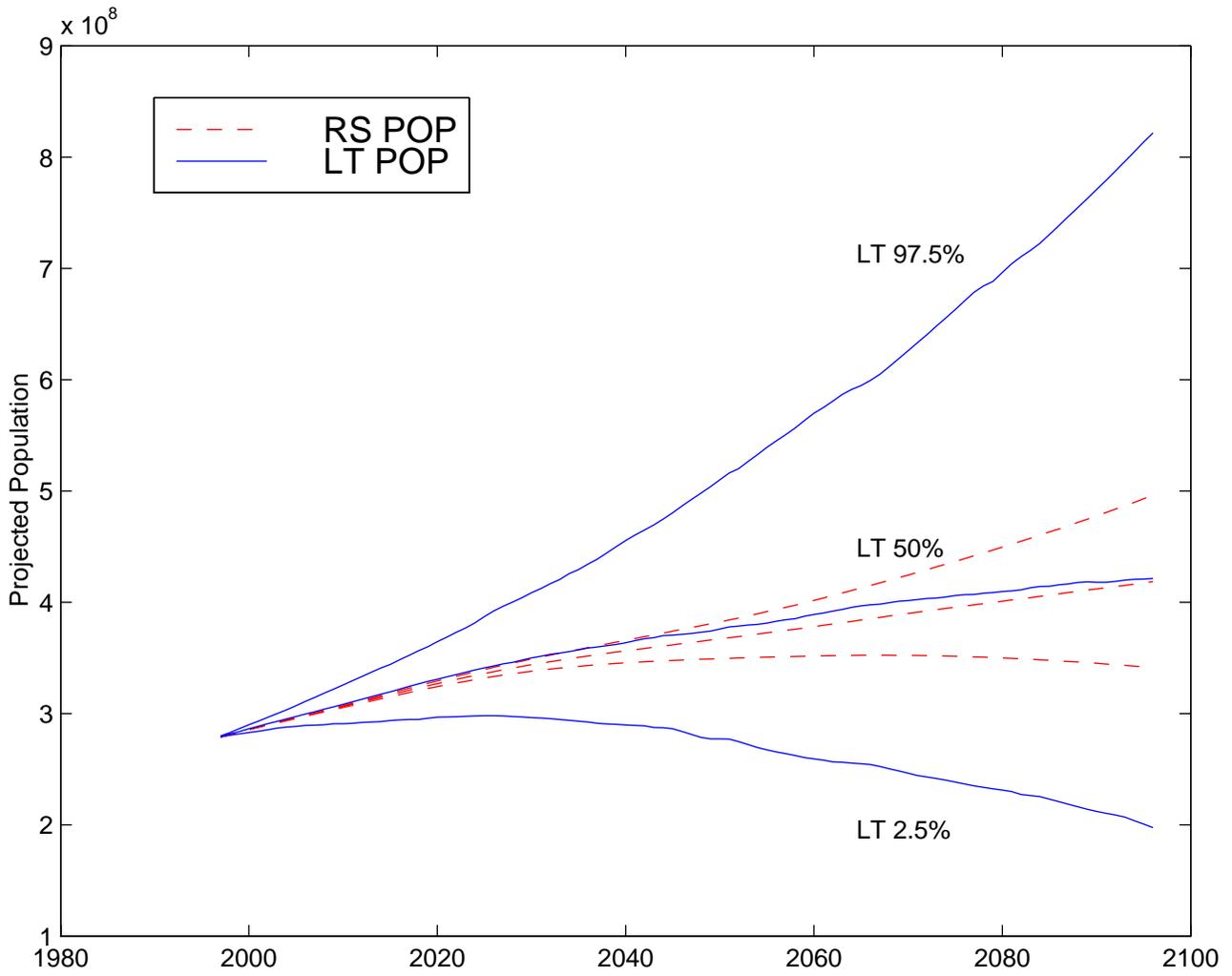


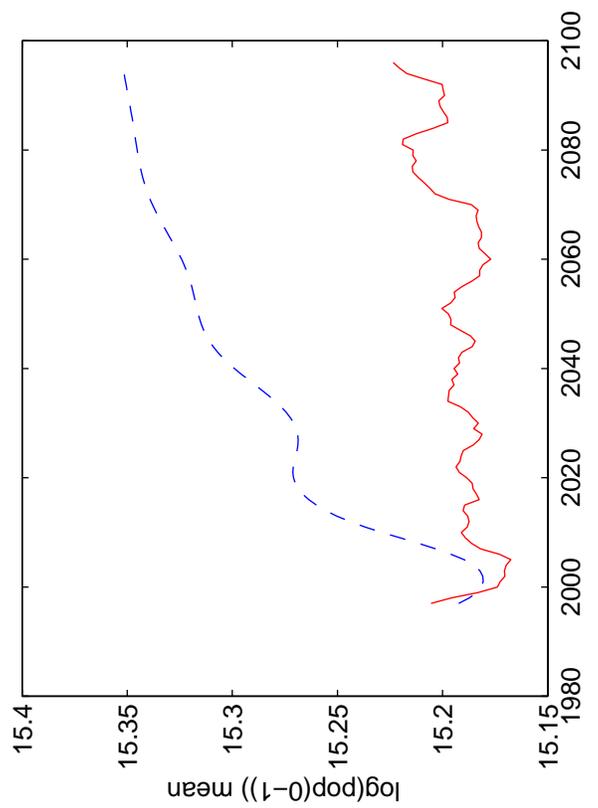
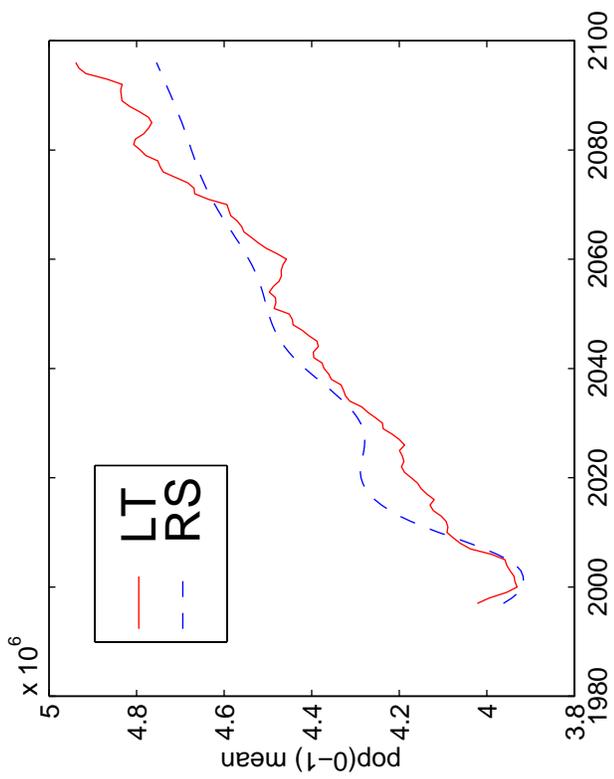
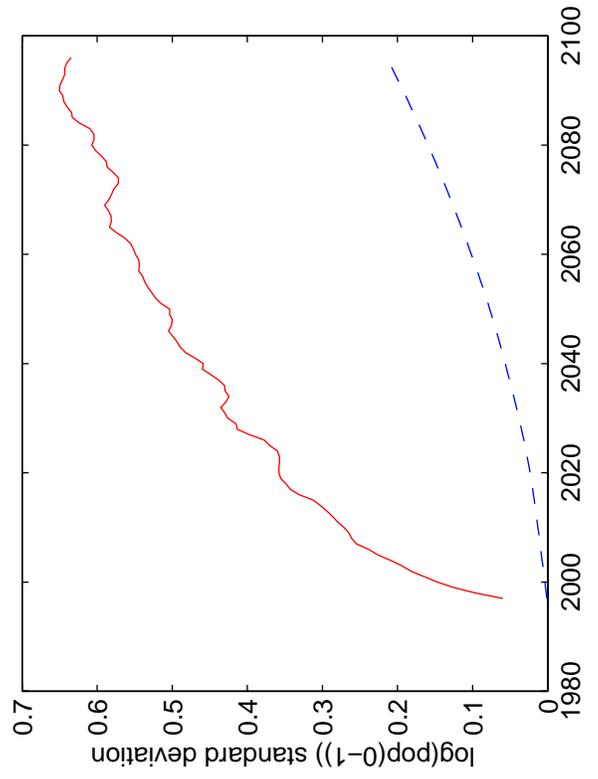
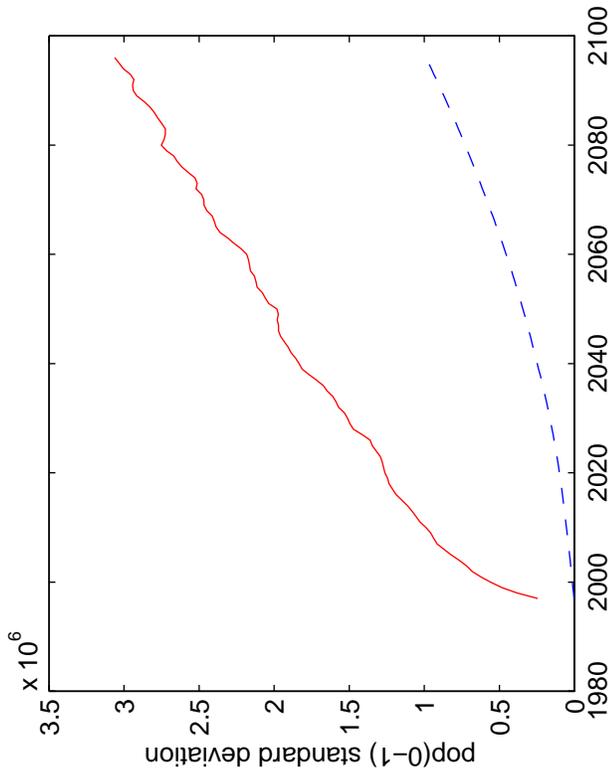


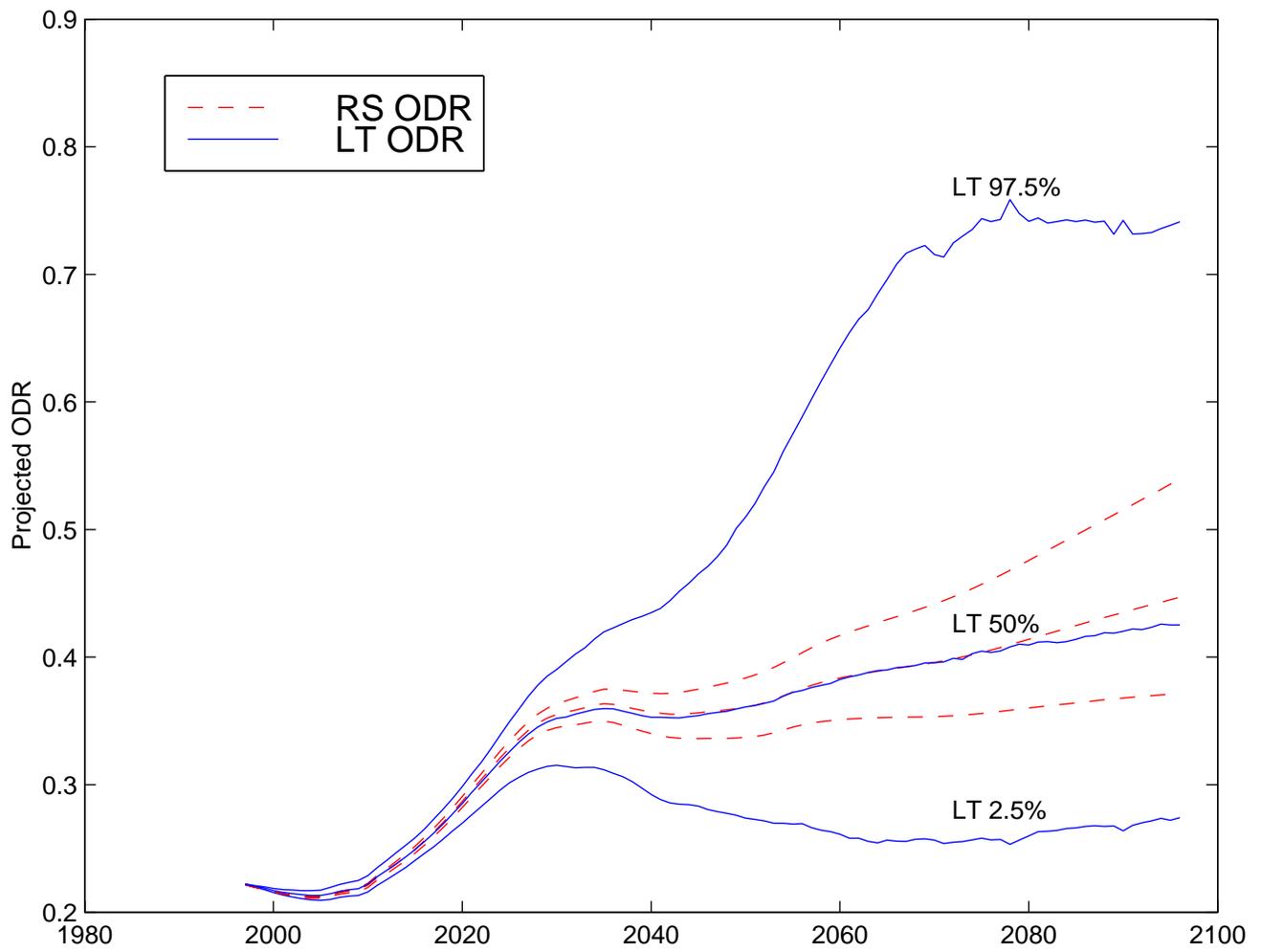


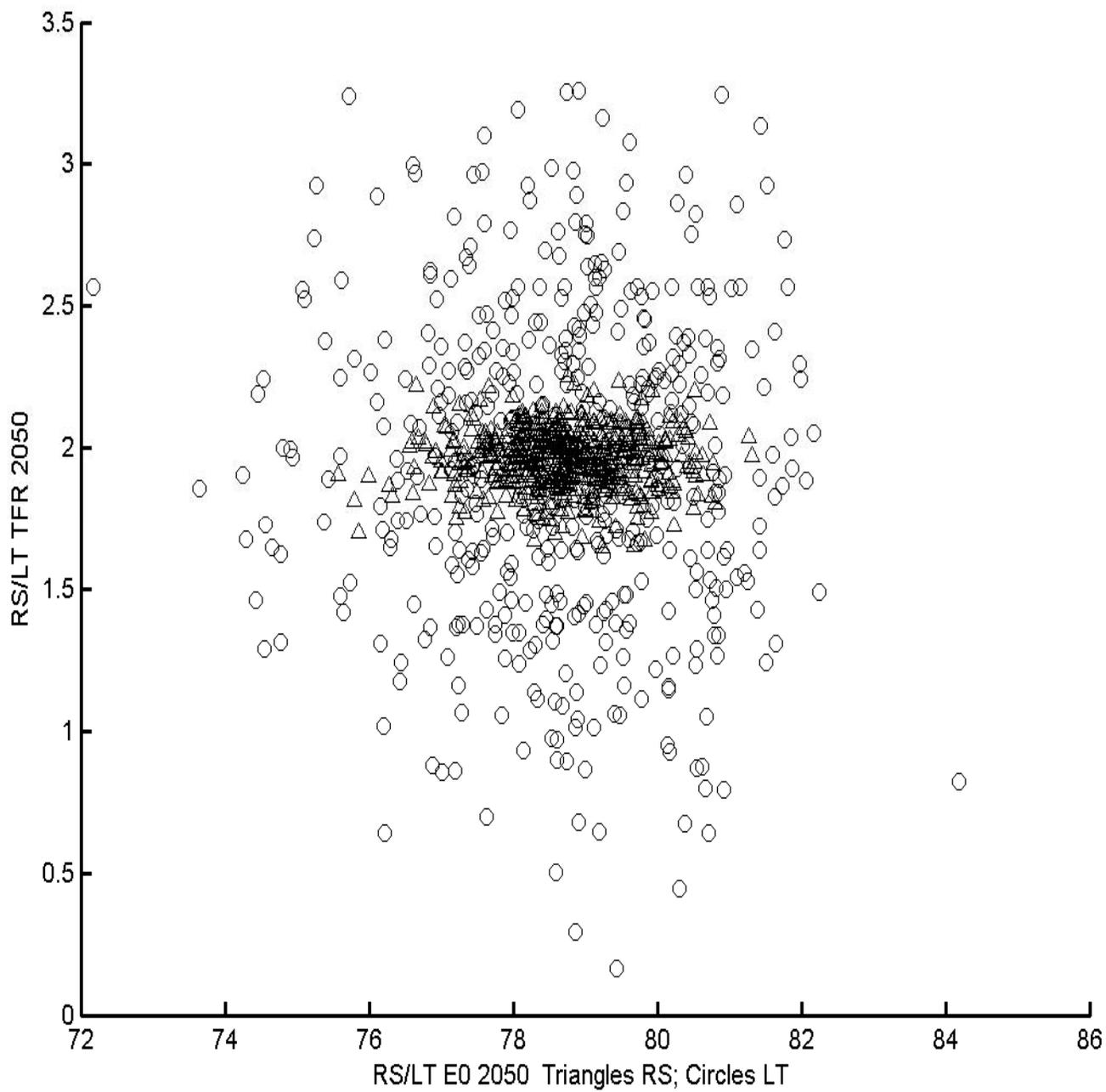
### SCALAR VARIANCE RS AND LT











### SERIAL AUTOCORRELATION OF OLD DEP. RATIO

