NBER WORKING PAPER SERIES

STOCHASTIC INFINITE HORIZON FORECASTS FOR SOCIAL SECURITY AND RELATED STUDIES

Ronald Lee
Timothy Miller
Michael Anderson

Working Paper 10917
http://www.nber.org/papers/w10917

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2004

This research was supported by the U.S. Social Security Administration through grant #10-P-98363-1 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA, any agency of the Federal Government, or the NBER. The authors also acknowledge support from Berkeley's NIA-funded Center for the Economics and Demography of Aging. The research funded here builds on basic research funded by NIA grant R37-AG11761. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Stochastic Infinite Horizon Forecasts for Social Security and Related Studies
Ronald Lee, Timothy Miller, and Michael Anderson
NBER Working Paper No. 10917
November 2004
JEL No. H0, H5

ABSTRACT

This paper consists of three reports on stochastic forecasting for Social Security, on infinite horizons, immigration, and structural time series models. 1) In our preferred stochastic immigration forecast, total net immigration drops from current levels down to about one million by 2020, then slowly rises to 1.2 million at the end of the century, with 95% probability bounds of 800,000 to 1.8 million at the century's end. Adding stochastic immigration makes little difference to the probability distribution of the old age dependency ratio. 2) We incorporate parameter uncertainty, stochastic trends, and uncertain ultimate levels in stochastic models of wage growth and fertility. These changes sometimes substantially affect the probability distributions of the individual input forecasts, but they make relatively little difference when embedded in the more fully stochastic Social Security projection. 3) Using a 500-year stochastic projection, we estimate an infinite horizon balance of -5.15% of payroll, compared to the -3.5% of the 2004 Trustees Report, probably reflecting different mortality projections. Our 95% probability interval bounds are -10.5 and -1.3%. Such forecasts, which reflect only "routine" uncertainty, have many problems but nonetheless seem worthwhile.

Ronald Lee
Demography and Economics
University fo California
2232 Piedmont Avenue
Berkeley, CA 94720
and NBER
rlee@demog.berkeley.edu

Timothy Miller
Center for the Economics and Demography of Aging
University fo California
2232 Piedmont Avenue
Berkeley, CA 94720
tmiller@demog.berkeley.edu

Michael Anderson
Center for the Economics and Demography of Aging
University fo California
2232 Piedmont Avenue
Berkeley, CA 94720
mikeand1@pacbell.net

This report consists of three sub-reports, which are listed below, followed by brief summaries of the conclusions of each.

Report I. A Probabilistic Forecast of Net Migration to the United States. (Miller and Lee)

Summary of Conclusions

1. Given the history of immigration to the US, a number of key assumptions must be made without a satisfactorily firm basis, such as whether to model numbers or rates, over what historical period to fit the model, and whether to include a trend in the forecast, or to impose central tendency based on expert opinion.
2. Experiments with a variety of approaches suggest that the probability distribution of the immigration forecast is not highly sensitive to these variations, although forecasts of the rate, with a trend, do lead to forecasts of higher numbers in the future.
3. Our preferred projection is based on numbers of immigrants rather than rates, and randomly samples trends between 0 and the historical average for each sample path. In this case, the projected number of net immigrants (legal and illegal) drops from current levels down to about one million in 2020, and then slowly rises to 1.2 million at the end of the century. The lower 2.5% probability bound is near 800,000 throughout the century, after the first decade or two. The upper 97.5% bound starts at 1.3 million, and rises quite linearly to 1.8 million at the end of the century.
4. With this range of models and forecasts, including immigration in the population forecasts makes little difference to the probability distribution of the old age dependency ratio, which is the item of prime importance for the Social Security forecasts.

Report II. Structural time series models and parameter uncertainty in Stochastic Projections of Social Security Finances. (Anderson and Lee)

Summary of Conclusions

We have experimented with a variety of different specifications of the time series models for wage growth and fertility, which are two of the key inputs for the projections. The expectation was that introducing parameter uncertainty, stochastically varying trends terms, and uncertain ultimate levels, would make the projections more uncertain. We did indeed find this to be so in every case, although one version of Homer's model, in which fertility was first logged, then modeled, then exponentiated, turned out to give a narrower probability interval than our other models including the standard ones. In some cases, the change in probability intervals for the individual input series was very slight, for example when parameter uncertainty was introduced to the fertility model, or when we used structural methods for wage growth. The big differences come from using an uncertain ultimate level for wage growth, or a structural estimate for fertility.

Although some of these new models have a substantial effect on the estimated probability distributions for the forecasts of the inputs themselves, they seem to make much less difference when they are embedded in a more fully stochastic Social Security projection. This is good news for the stochastic projections, because it suggests that they are not so sensitive to the specifications of the input series as one might have feared. This is true in our stochastic forecasting model which has only four stochastic inputs. It would be even more true in the forecasting models of Social Security and CBO with their greater number of inputs. One would not want to push this argument too far, of course. Ultimately, the stochastic forecasts of Social Security are only as good as the stochastic forecasts of the key input series.

Report III. Stochastic Infinite Horizon Forecasts of Social Security Sustainability. (Lee and Anderson).

Summary of Conclusions

1) Many issues surround infinite horizon forecasts, and the whole enterprise can certainly be questioned. Nonetheless, we have found it useful simply to extend the range of the stochastic forecasting models to very distant horizons. We call these "routine" or "business as usual" stochastic forecasts, because their uncertainty does not reflect the possibility of structural shifts. They understate actual uncertainty.
2) Both the Flat Fund Ratio Tax measure and the Unstable measure are useful simple approximations to the deterministic or median infinite horizon open group imbalance measure. The Flat Fund Ratio is the immediate and permanent tax increase that would be needed to hold the ratio of the Trust Fund to Costs constant over the last two years of the 75-year projection. It is more intuitive and therefore easier to explain than the Unstable measure, but it underestimates the imbalance, whereas the Unstable measure (explained in the report) gives a very good approximation to the infinite horizon measure, at least under current circumstances.
3) The 2004 Trustees Report indicates an infinite horizon open group imbalance equal to 3.5% of payroll, consistent with Lee and Yamagata's (2003) calculation using SSA mortality assumptions. Based on our 500-year projection with our own mortality forecasts, we estimate it to be 5.15%, substantially larger. Our two simple methods, based on our 75 year projections, indicate levels of 4.36% for the Flat Fund Ratio measure, and 5.21% for the Unstable measure.
4) Good estimates of the uncertainty of the simple measures cannot be derived from stochastic forecasts over the 75 year horizon, at least using the methods we have attempted. Therefore the simple measures are useful only for central tendency.
5) The "routine" uncertainty surrounding the infinite horizon estimates of Summary Actuarial Balance is about 40% greater than the uncertainty of the 75 year projections: the 95% probability interval is 9% wide versus 6.5% for the 75-year horizon.
6) Raising tax rates immediately by an amount intended to achieve sustainability would imply substantial chances of huge Trust Fund accumulations that neither could nor should be realized in practice, at least not through holdings of government bonds. Adaptive policies that maintain the Trust Fund ratio at a desirable level seem more attractive, but have not yet been explored.

Report I

A Probabilistic Forecast of Net Migration to the United States

Timothy Miller
and
Ronald Lee

Center for the Economics and Demography of Aging
University of California at Berkeley
2232 Piedmont Avenue
Berkeley, CA 94720-2120
tmiller@demog.berkeley.edu

October 8, 2004

The research reported herein was performed pursuant to a grant from the US Social Security Administration (SSA) as part of the Retirement Research Consortium. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA or any agency of the Federal Government. The authors also acknowledge support from Berkeley's NIA-funded Center for the Economics and Demography of Aging. The research funded here builds on basic research funded by NIA grant R37-AG11761.

Background

The core engine of stochastic macro forecasting models for Social Security finances is a stochastic population model. The first such model was developed by Lee and Tuljapurkar (1994), henceforth LT94. This model treated fertility and mortality as stochastic, but took immigration as deterministically given at some specified rate. There were two reasons for doing this. One was that immigration could be viewed as a policy variable. A policy maker would not find it useful to be given a forecast in which a key policy level was treated as stochastic, and outside control. Another is that immigration appears to be more difficulty to forecast then either fertility or mortality, because its trend has been so problematic in the US. In this report we discuss the issues and options for forecasting immigration stochastically, and develop some different versions of stochastic forecasts of immigration. These are then incorporated in a LT style model for generating stochastic population forecasts, and the results are discussed.

Forecasts of net migration to the United States are issued by the U.S. Social Security Administration (SSA), the U.S. Census Bureau (CB), and the United Nations (UN). The different approaches of these agencies are discussed in the following paragraphs and summarized in Table 1.

Social Security Administration

SSA forecasts immigration as a separate component of their population forecast. The projection assumes that changes in population size, age structure, and composition have no impact on the annual flow of net immigrants. The SSA median forecast assumes no change in current law in which approximately 600,000 legal immigrants are admitted each year. Illegal immigration is assumed to continue at half the level of legal immigrants. In the short-run, owing in part of the effects of the recent amnesty, SSA forecasts 800,000 legal immigrants plus another 400,000 illegal immigrants for a total of 1.2 net immigrants per year. They project a decline in the total number of net immigrants toward a long-run average of 900,000 per year (600,000 legal plus 300,000 illegal). Their high cost variant envisions net immigration of 1.3 million per year and while the low cost variant sees a decline to 672,500 net immigrants per year. Recently, the Social Security Technical Panel on Assumptions and Methods (2003) recommended that the SSA Trustees change their forecast methodology and project using a net immigration rate, rather than numbers. We explore the implications of both methods in our projections discussed in the next section.

Census Bureau

The Census Bureau forecasts in-migrants separately from out-migrants. The number of out-migrants is based on an out-migration rate applied to the projected foreign-born population. In the baseline projection, the number of in-migrants in the short-run (thru 2020) is based on assumptions about trends in immigration by category: immediate relatives, numerically-limited categories, refugees, and illegal immigrants. From 2021 to

2030, the projection assumes the number of in-migrants rises in response to increases in the dependency ratio. From 2030 onward, the number of in-migrants is assumed to be fixed at the new limit of 1.45 million immigrants. While the number of in-migrants is constant since 2030, the number of out-migrants fluctuates over the period in response to shifts in the age composition of foreign-born. Net immigration reaches its peak in 2030 at 1.06 million and slowly decreases to 926,000 by 2100. The long-run values of the baseline forecasts of SSA and CB are surprisingly close. In contrast, the high-low range of the Census Bureau is very large relative to SSA. By 2100, the CB high-low estimates differ by 3 million immigrants (3 million versus 113,000). This is about 6 times as wide as the SSA high-low estimates which differ by only half a million (1.3 million versus 672,000). Recently, the CB issued a new interim population forecast (CB, 2004). In this forecast their immigration assumptions is set equal 0.938 times the year 2000 middle forecast plus 0.62 times the year 2000 high forecast. This implies immigration will reach nearly 1.1 million per year by 2100.

United Nations

The UN forecasts net immigrants for each country based on the policy stance of the country. In the case of their U.S., they forecast a decline from 1.25 million net immigrants per year in the late 1990s to 1.1 million by 2010, followed by a constant level of 1.1 million. The UN forecast is quite close to the middle forecast of both CB and SSA. The UN provides no alternative migration scenarios. Table 1 below summarizes the differences among the SSA, CB, and UN in their immigration forecasts.

Table 1. Summary of Methods for Immigration Forecasts by the Social Security Administration, Census Bureau, and United Nations.

	SSA	CB	UN
What is forecast?	Net immigrants.	In-migrants separately from out-migrants.	Net immigrants
Alternative migration scenarios?	High and low cost.	High and low population sizes.	None.
Baseline forecast is based on:	Current law with temporary high net migration due to amnesty.	Slight increase in response to increased in OADR in the decade of the 2020s.	Policy stance of county.
Baseline numbers:	1,200,000 (of whom 400,000 are illegal) declining to long-run average of 900,000 (of whom 300,000 are illegal) in 2024.	964,000 net immigrants (1,251,000 in-migrants and 287,000 out-migrants). Declines to 713,000 net migrants in 2010, then peaks at 1,061,000 net migrants in 2030, slowly declines reaching 926,000 net migrants in 2100. In 2004, CB issued new interim projections which raised the median immigration forecast to equal 0.938 times the old median forecast plus 0.062 times the old high forecast. This implies a 2100 forecast of nearly 1.1 million net immigrants.	1,250,000 declining to long-run average of 1,100,000 in 2010.
Alternative scenarios:	Low cost: 1,300,000 (of whom 450,000 are illegal). High cost: 672,500 (of whom 200,000 are illegal)	Low pop: 624,000 net migrants in 2000, declining to 113,000 net migrants in 2100. High pop: 1,363,000 net migrants in 2000, increasing to 3,047,000 in 2100.	No alternative migration scenarios (only fertility scenarios).
Issued:	2004 Trustees Report	2000 National Population Projections. Since replaced with the 2004 National Population Projections with higher immigration levels.	World Population Prospects: The 2000 Revision (2001).

The Time Series Approach to Forecasting Net Immigration

We develop a probabilistic forecast of immigration. Rather than examining three scenarios as done by SSA or CB, we develop thousands of possible sample paths based on both time series analysis of past variation and on expert opinion about the future course of immigration. There are several distinct advantages of modeling uncertainty based on an analysis of historical times series rather than the use of high-low scenarios. (For an extensive discussion see Lee (2004) and Lee and Tuljapurkar, (2000)). The time-series approach deals with uncertainty in a probabilistic and consistent fashion. By contrast, in the scenario approach, the high-low range generally lacks any probabilistic interpretation, so the user has no sense of how likely they are to contain the true values. The scenario approach assumes perfect correlation of the component trajectories (e.g., fertility, mortality, and migration) with each other and across time. For example, high fertility is always high and is always coupled with high immigration. It is generally not clear whether the high-low range is meant to contain annual variations or long-run averages. The probability bands for long-run averages will be considerably narrower than bands for annual variations due to error cancellation over time.

Ideally, we would be able to estimate a plausible and appealing time series model of the stochastic process based entirely on historical data, which would then be sufficient to identify both the long run level or trend in the input series for forecasting purposes, as well as identifying the variances and covariances of interest. In the case of mortality forecasts, this ideal case holds quite well, using the Lee-Carter (1992). However, in the case of fertility, there are strong theoretical reasons to expect the central tendency to have changed over the course of the 20th century, but not necessarily to continue to change in the future. Also, there is external information about the range of possible variation that should be taken into account (Lee, 1993). For fertility, then, the ideal conditions are not met, and the forecast model takes a long term expected value from expert opinion, and excludes certain stochastic realizations as impossible. The real interest rate and the rate of growth of covered wages are handled in a similar manner, so far as central tendencies are concerned. For immigration, it seems clear that the historical data does not provide a fully adequate basis for assessing the long term trend. Even more than fertility, the course of immigration in the 20th century US was dominated by very long trend like movements, which might or might not reasonably be expected to continue in the future. Thus there are difficult strategic issues to be confronted before we begin modeling, as we now consider. These issues are not specific to probabilistic modeling, but rather are the same issues that anyone constructing a deterministic projection would also have to confront.

Here are some of the basic questions:

- Should we forecast the rate of immigration or the number of immigrants?
- Should we rely on the observed historical trend for our forecast?
- Should we use expert opinion to specify the likely trend?
- If relying on historical data, how far back in history should we go?

The National Research Council (2000) published a report assessing the projections methodology of the UN, the World Bank, and the International Division of the US

Census Bureau. This report evaluated methods used for global projections, covering all countries of the world, which is not the same as preparing a projection for the US. Nonetheless, the chapter on immigration in the report does provide some useful insights for present purposes. Here are some of the relevant conclusions.

1) 1) Errors in projecting immigration lead to errors in population projections that are as important as those arising from errors in projecting fertility, and far more important than errors in projecting mortality, in the United Nations projections record since 1960 or so. However, the most important errors arise for countries that are subject to a demographic “quake”, an unexpected political, military or environmental shock. Excluding these instances, the role of errors in forecasting immigration is only about half as great, comparable to mortality. We would not expect that “quakes” of this sort would be likely for the US.
2) It is useful to distinguish countries that have a long history as receiving countries for immigrants, since they are much more likely to continue in this role in the future. The US is a prime example. However, it is also important to realize that some long-time receiving countries such as Brazil and Argentina no longer receive many international immigrants, so reversals do occur and continuing status as a receiving nation cannot be taken for granted but must be assessed.
3) Although it is plausible that increasing globalization of markets will lead to increased international migratory flows in the coming decades, it is also possible that national policy will tend toward restricting immigration. It is also possible that other industrial nations will increasingly compete with the US for immigrants from the Third World as an antidote to population aging, as is much discussed these days.
4) Using UN data since 1950, experiments with different projection assumptions for immigration and emigration were carried out. For our purposes, the two most interesting possibilities were considered were a) assuming 0 net immigration in the future regardless of past levels or rates, and b) assuming that the most recent level of immigration (over a five year interval) continued constant in perpetuity. Forecasting errors were then assessed. It was found that for horizons up to ten years, the constant immigration flow assumption led to smaller errors than the zero assumption, but for horizons of 15 to 40 years, the constant immigration assumption led to larger errors. This was true for all countries considered together, and also true if countries experiencing “quakes” were considered separately.

The National Research Council (2000) report does not explicitly address any of the questions posed earlier, but it does provide some general ideas about how to proceed. In the US, we have had a fifty year history of constantly accelerating immigration. The report can be read as suggesting some caution in assuming that this long and impressive trend will continue.

Forecasting the Rate or Level of Immigration?

Typically, forecasts are made of the *number* of immigrants rather than the immigration *rate*. Forecasting numbers of immigrants is consistent with legislative-controls on immigration which focus on numbers of legal permanent residents admitted each year. For example, under current U.S. immigration law the number of legal permanent residents admitted each year is limited to 675,000 plus unlimited numbers of immediate relatives (spouses and unmarried children U.S. citizens and parents of adult U.S. citizens) and unlimited numbers of previously-admitted refugees and asylees who are adjusting their immigration status. For those admissions subject to the annual numerical cap, it is easy to observe the effects of these limits in the 5.3 million immigrant applications that are currently pending or in the average 10 year waiting time for admission of siblings of naturalized U.S. citizens.

Recently, the SSA Technical Advisory Panel recommended that the Trustees long-run target for intermediate forecast be based on a net migration rate rather than a net migration level. The Panel recommended the long-run target be set to a net migration rate of 3.2 per 1,000 (based on Wilmoth's estimate of the net migration rate for the period 1821-2002). The panel recommended that the high-cost scenario be based on the continuation of the current level of immigration (implying a declining net migration rate) and that the low-cost scenario be based on a continuation of the current net immigration rate (4.2 per 1000). The panel listed several factors which might lead to more restrictive immigration: economic slowdown, national security concerns, falling fertility in sending regions, the dissipation of the residual immigration effects of the IRCA legalization of the 1990s, and the increased demand for immigrants in rapidly aging Europe and Japan. They also listed several factors which might lead to increased immigration: demand for immigrant labor is likely to scale with the growth of population and the economy, the increased share of foreign born increases the network available for immigrants, there are large untapped pools of immigrants around the world, and the difficulty in immigration enforcement. On balance they argued: "Faced with these varied arguments, the Panel concluded that there is no strong reason to anticipate a sharp break with past trends in the near future" (Technical Panel 2003).

There is another point to consider when deciding to forecast rates versus levels. In the context of forecasting rates stochastically, we might ask whether we expect a positive or negative co-variation between net immigration rates and population size. On the one hand, we might expect positive co-variance arising as larger populations result from higher immigration rates. On the other hand, we might expect negative co-variance as slower population growth might lead to increased demand for immigration. This ambiguity is one reason why projecting the number of immigrants might be preferable, because the historical co-variations of the rate of immigration and the base population size to which it is applied are implicit in the projection of their product, the numbers of immigrants.

In the models and forecasts which follow, we analyze four combinations of assumptions: forecasts of the net immigration rate or level and use of the observed historical trend in the forecast or substitution of a constant rate or level in the forecast. In addition, we present a forecast in which we assume that the trend in levels of immigration are uniformly distributed over the interval between 0-trend and the time series estimate of the trend.

Choice of a time period

We begin by modeling stochastic immigration in a similar manner to the stochastic forecasts of fertility as in Lee (1993). We issue a probabilistic forecast of the *level* of net immigration based on a time-series analysis of the historical series of legal immigration to the United States. A similar method is used to generate a probabilistic forecast of the net immigration *rate*. The first question that arises in the context of immigration is the appropriate time period for the time-series estimation.

Figure 1.1 shows the gross legal immigration rate from 1820 to 2002. Two distinct periods are evident in the series. The 19th century is characterized by high and fluctuating rates, while the 20th century shows much less variation with a distinct upward trend. For the time series estimation we chose the period following the Federal Immigration Act of 1924 which permanently established immigration limits via the quota system. Choice of a shorter or longer time-series would influence both the predicted variation and trend of the forecast.

The distinct peak in legal immigration in the early 1990s represents an immigrant amnesty. We chose not to model this in the simulation, since our forecast is most concerned with when people entered the country and not when they were legalized. So, rather than include amnesties as a periodic occurrence, we stochastically forecast legal immigration and assume that illegal immigrants represent a fixed proportion of legal entries. We use the SSA assumptions that illegal immigrant flows are ½ that of legal flows and that ¼ of immigrants eventually emigrate. So, our forecast of legal immigrants is multiplied by 1.125 (=3/2*3/4) to transform it into a forecast of net immigrants (legal plus illegal).

Use of the historical trend

The second question that arises is whether to allow the observed historical trend to influence our probabilistic forecasts or to replace this with a trend provided by expert opinion. On the one hand, previous studies of both mortality and fertility forecasts have found expert judgment to be unduly influenced by the recent past. On the other hand, if we have reason to believe there has been a structural shift or when there is doubt about the sustainability of current trends, then subjective expert opinion would be preferred. In the case of probabilistic forecast of fertility (Lee 1993), it was argued that the fertility transition had resulted in a structural shift. Therefore, the wiser course of action was to impose a lower level of fertility than that observed in the historical series. This was done via a mean constrained forecast such that the mean across all trajectories tended toward the long-run average deemed to be 1.9 births per woman.

Figure 1.2 presents the *level* of gross legal immigration to the U.S. from 1820 to 2002. This series is similar to that of the gross legal immigration rate seen in Figure 1.1. The 20th century is characterized by a distinct upward trend in both the level and rate of immigration.

Figure 1.3 presents data from the United Nations on net immigration flows from developing nations to developed ones from 1950-2000, along with the UN forecast for 2000-2500. The latter part of the 20th century is characterized by a distinct upward trend in the net immigration rate to developed nations. Yet, as is evident here, the UN forecasts a distinct break with the past trend. It would be interesting to investigate whether previous UN forecasts also envisioned a distinct break which was subsequently shown to be inaccurate. The UN, SSA, and CB all forecast a leveling off of the immigration to the US – in sharp contrast to the experience of recent decades.

In our probabilistic simulations of immigration, we consider alternative forecasts in which the historical trend is replaced with a subjective assumption of no trend which is consistent with the opinions expressed by the UN, SSA, and CB.

Fitting the Time Series Model

Lacking data on flows of illegal immigrants and out-migrants, the time-series was fit to data on legal immigrants admitted to the U.S. from 1925-2002. Illegal immigrants who were granted legal permanent residence under the amnesty program of the 1990s (IRCA) were not included in this series. Auto-regressive and moving average time series models with a linear trend (in order to de-trend the series) were fit to the level of immigrants and the rate of legal immigrants. Based on Akaike information criterion, the AR(2) model with linear trend was found to have the best fit for both series. (Results in Tables 2 and 3). We also attempted to de-trend the data by taking first differences. However, in this case, the probability bounds of the resulting forecast widen very rapidly so as to reach quite extreme levels of immigration within a decade. We don't have a good justification for assuming the series is trend-stationary rather than difference-stationary, beyond the implications observed in the forecasts.

Table 2. Time-series estimate of Legal Immigrants to U.S. 1925-2002

Coefficients:

	ar1	ar2	intercept	seq(1925, 2002)
	1.0077	-0.1373	-21958445	11420.363
s.e.	0.1114	0.1232	4384448	2230.289

sigma^2 estimated as 4.995e+09: log likelihood = -982.4, aic = 1974.8

Table 3. Time-series estimate of Legal Immigration Rate to U.S. 1925-2002

Coefficients:

	ar1	ar2	intercept	seq(1925, 2002)
	1.0599	-0.1721	-50.4855	0.0268
s.e.	0.1107	0.1189	24.6147	0.0125

sigma^2 estimated as 0.1163: log likelihood = -27.66, aic = 65.31

The point estimates of the AR coefficients imply a stationary forecast. However, in generating immigration trajectories, we sample from the joint distributions of the two AR coefficients. Therefore, it is possible that for some of our samples the AR coefficients imply a non-stationary path. we could have restricted the sampling so that such combinations were rejected. Instead, we have placed bounds on the resulting forecasts. In the forecasts of immigration rates, we established an upper bound of +15 per 1,000 (based on US historical experience) and a lower bound of -15 per 1000 (based on post-war experience in more developed countries). In the forecast of immigration level, we establish an upper bound of 4 million net immigrants per year.

Previous Work

In previous work, we stochastically forecast domestic and international migration to California for a probabilistic population forecast of the state (Miller, 2002). Net domestic immigration to California was based on a time series analysis of the net immigration *rate*. The rate was then applied to the population total in the previous year in order to derive a level of net immigration. The age/sex combination distribution of domestic immigrants was taken from the March CPS and was assumed to remain unchanged over the forecast. A similar approach has been taken in this paper. International migration to California was based on a time series analysis of the proportion of immigrants intending to reside in California. The forecast of this proportion was then multiplied by the Middle Series Forecast of Immigration to the U.S. from the Census Bureau in order to obtain the number of international migrants to California.

Results

Figure 1.4 compares our immigration forecasts to those of the UN, SSA, and CB. Our stochastic forecast of the level of immigration without the historical trend yields a median forecast nearly identical to that of Social Security. Our 95% probability interval is about the same width as the high-low bounds of Social Security, but both our upper and lower bounds on the 95% probability interval lie above that of Social Security. The high-low bounds by Census Bureau are more than 4 times as wide as our 95% probability interval.

Figure 1.5 shows our stochastic forecast of the level of immigration using the historical trend present in the time series. In this case, the baseline scenarios of both SSA and CB lie below our 95% probability interval. This reflects the fact that the baseline forecast of both agencies represents a distinct break with the past. The probability distribution for the cumulative average of the forecast looks odd in this figure. We would expect the distribution for the cumulative average to lie within the distribution for the forecasts for individual years, but this only need happen when there is no trend in the central forecast. When there is a rising trend, as in this version of the model, the cumulative average forecasts for longer horizons reflect the lower values of forecasts at shorter horizons, which pull down the whole distribution below the individual year forecasts.

The next set of figures compares the results of our stochastic forecast of the immigration *rate* to the forecasts of Census, SSA, and UN. A distinct narrowing of the high-low bounds of Census and SSA are evident in Figure 1.6. This is a by-product of the assumption of constant immigration levels with a growing population. In contrast, our probability intervals (based on rates) expand over time. By 2100, our probability intervals for the cumulative immigration rate are about as wide as those of the CB, and many times wider than those of SSA.

Figure 1.7 shows the result of our stochastic forecast of the immigration rate using the trend present in the time series. In this case, our median forecast shows a slight increase in the net immigration rate from the current level of 4.2, reaching 4.7 by 2100. As the forecast horizon increases, the forecasts of CB, SSA, and UN fall below our 95% probability interval. Only CB's high population scenario lies within our 95% probability interval at the end of the forecast.

In our view, the biggest uncertainty is in the choice of trend. We do not see a compelling reason to prefer either the trend extrapolation of the time series trend estimate or the expert opinion (expert opinion points in the direction of ignoring past trends). Therefore we created a mixed model in which we assume that the trend for an entire sample path is distributed uniformly between zero and the time series estimate of trend, for numbers of immigrants. In addition, there is the usual uncertainty due to innovations in the model. The results of this approach are plotted in the remaining diagrams with the label "uncertain trend". These are our preferred results. They generally lie midway between the forecasts with trend and without trend, for both central tendency and probability interval.

We will now consider the effects of modeling immigration stochastically rather than deterministically. We have used a stochastic population model we developed for use in an analysis of a probabilistic forecast of Medicare (Lee and Miller, 2002), which is similar to Lee and Tuljapurkar (1994). Figure 1.8 compares our alternative stochastic immigration forecasts with a deterministic immigration forecast. In terms of projected population size, there is little difference between the forecast based on stochastic *level* of immigration without a trend and the forecast based on deterministic immigration. Partly, this reflects the fact that we already have a lot of variance in the model from the two other stochastic components (TFR and e(0)). In particular, forecast of the *level* of immigration adds relatively little variation to the forecast relative to that of the TFR. The 95% probability interval for the cumulative value of the TFR is about +/- 25% of the baseline TFR. In 2002, there were 4 million births in the US, a variation in TFR of +/- 25% would translate into +/- 1 million additional births. In contrast, our probability interval for immigration based on a model of immigration level without time series trend leads to differences of +/- 350,000 immigrants. This is about 1/3 as large an effect of variation as in the TFR. In addition, migrants contribute less to population size than do births, since the average age of migrants is near 30. So, the actual effect of an immigrant on population size might be only 5/8 as large as a birth. The variation in population size introduced by modeling immigration level must be closer to 1/5 as large an effect as modeling variation in TFR.

Figure 1.8 shows much larger upside uncertainty when rates of immigration are forecast rather than levels, because in this case upside uncertainty in rates interacts multiplicatively with upside uncertainty in population size, which does not happen when numbers of immigrants are directly projected. We are skeptical of these large upside probability ranges, because it seems likely to us that rates might be negatively correlated with population size, whereas our projections assume these to be independent. However, much of this is speculation and there are no compelling arguments that we see on either side.

Figure 1.9 shows the results of three stochastic forecasts: one in which only fertility and mortality are stochastically forecast, the second in which only immigration is stochastically forecast, and the third in which all 3 components are stochastically forecast. Here we see that the probability interval based on a stochastic forecast of immigration alone is about 1/4 as wide as that of simulation based on stochastic fertility and mortality which is close to our admittedly crude approximation of 1/5. In contrast to these results, the stochastic forecasts based on *rates* show wider probability intervals as these forecasts allow for more extreme population scenarios in which a high level of population can be matched with a high immigration rate and vice versa. As expected, the median of the forecasts based on rates are higher than those forecasts based on level of immigration.

Turning to a consideration of the old-age dependency rate in Figure 1.10, we find that forecasts based on immigration rate tend to produce younger populations than those based on levels. We also note that over the first 50 years of the forecast (see Table 4 below), the OADR in all these methods are surprisingly close – suggesting that uncertainty in immigration would contribute little to uncertainty in social security finances.

Table 4. OADR in 2050. Simulation begins in 2003 with OADR at 216 per 1,000.

	Type of immigration forecast:				
	Deterministic	Level no trend	Level with trend	Rate no trend	Rate with trend
2.5 percentile	312	310	304	299	288
50^{th} percentile	382	382	370	372	363
97.5 percentile	472	467	458	461	456

Foreign born population

In a stationary population, we can calculate the percent of the population which is foreign-born based on the person-years lived of natives versus foreign-born. Assume that there is no difference in mortality between natives and foreign-born. In this case, the difference between the person-years lived in the U.S. by immigrants and by natives is simply (T(0)-T(30))/T(0)) where 30 is taken to be the average age of arrival of an immigrant. Using a recent life table, we can calculate this value as about 5/8. So the steady state estimate of the percent foreign-born in the population is simply (5/8*I)/(B + (5/8*I), where I is the net immigration rate and B is the crude birth rate Currently, the US the net immigration rate is about 4/1000 and the crude birth rate is about 14/1000. A steady-state calculation leads to a population of about 15% foreign-born.

Using Census data and the assumptions of the SSA forecasts, we have calculated the foreign born population as a percent of the working-age population and of the total population in Figure 1.11. We see that our crude estimate of 15% foreign-born from the preceding calculation is quite close to the projection for the US.

Projected Net Immigrants

Figure 12 shows the projected numbers and probability intervals of net immigrants, legal and illegal combined. Our preferred projection is based on the “uncertain trend” approach. The central forecast for this case shows the number of net immigrants dropping from current levels down to about one million in 2020, and then slowly rising to about 1.2 million at the end of the century. The lower 2.5% probability bound is near 800,000 throughout the century, after the first decade or two. The upper 97.5% bound starts at 1.3 million, and rises quite linearly to 1.8 million at the end of the century.

Summary and Conclusions

1. Given the history of immigration to the US, a number of key assumptions must be made without a satisfactorily firm basis, such as whether to model numbers or rates, over what historical period to fit the model, and whether to include a trend in the forecast, or to impose central tendency based on expert opinion.

2. Experiments with a variety of approaches suggest that the probability distribution of the immigration forecast is not highly sensitive to these variations, although forecasts of the rate, with a trend, do lead to forecasts of higher numbers in the future.
3. Our preferred projection is based on numbers of immigrants rather than rates, and randomly samples trends between 0 and the historical average for each sample path. In this case, the projected number of net immigrants (legal and illegal) drops from current levels down to about one million in 2020, and then slowly rises to 1.2 million at the end of the century. The lower 2.5% probability bound is near 800,000 throughout the century, after the first decade or two. The upper 97.5% bound starts at 1.3 million, and rises quite linearly to 1.8 million at the end of the century.
4. With this range of models and forecasts, including immigration in the population forecasts makes little difference to the probability distribution of the old age dependency ratio, which is the item of prime importance for the Social Security forecasts.

References

Burdick, Clark and Joyce Manchester (2003) "Stochastic Models of the Social Security Trust Funds" Office of Policy Office of Research, Evaluation, and Statistics, *Research and Statistics Note No. 2003-01* (March)

Hollmann, Frederick W; Tammany J. Mulder and Jeffrey E. Kallan (2002). *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100.* Population Projections Branch, Population Division, Bureau of the Census. Population Division Working Paper No. 38. Available online: *http://www.census.gov/population/www/documentation/twps0038.html*

Lee, Ronald (1993). "Modeling and Forecasting the Time Series of US Fertility: Age Patterns, Range, and Ultimate Level," *International Journal of Forecasting* v.9, pp.187-202.

Lee, Ronald (2004). *Quantifying Our Ignorance: Stochastic Forecasts of Population and Public Budgets.* Paper prepared for the Rand Summer Institute Gala Celebration for the NIA Centers for the Demography and Economics of Aging. Available online: http://repositories.cdlib.org/iber/ceda/papers/2004-0001CL/

Lee, Ronald and Shripad Tuljapurkar (2000). "Population Forecasting for Fiscal Planning: Issues and Innovations" in Alan Auerbach and Ronald Lee, eds., *Demography and Fiscal Policy*, Cambridge: Cambridge University Press. Available online: http://repositories.cdlib.org/iber/ceda/papers/2000-0004CL/

Lee, Ronald and Tim Miller (2002). "An Approach to Forecasting Health Expenditures, with Application to the US Medicare System," *Health Services Research*, v. 37, n. 5, pp. 1365-1386 (October 2002).

Miller, Tim (2002). California's Uncertain Population Future. Technical Appendix. Available online: http://www.demog.berkeley.edu/~tmiller/papers/p2002.california.forecast.pdf

National Research Council (2000) *Beyond Six Billion: Forecasting the World's Population*. Panel on Population Projections. John Bongaarts and Rodolfo Bulatao, eds. Committee on Population, Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

Social Security Administration, Board of Trustees (2004). *The 2004 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Available online: *http://www.ssa.gov/OACT/TR/TR04/index.html*

Technical Panel on Assumptions and Methods (2003). *Report to the Social Security Advisory Board*. Washington, D.C., October 2003. Available online: *http://www.ssab.gov/NEW/documents/2003TechnicalPanelRept.pdf*

United Nations, Population Division (2003). *World Population Prospects: The 2002 Revision.* Database available online: http://esa.un.org/unpp/

Figure 1. Legal Immigration Rate: US, 1820−2002
(Dashed spike indicates amnesty immigration)

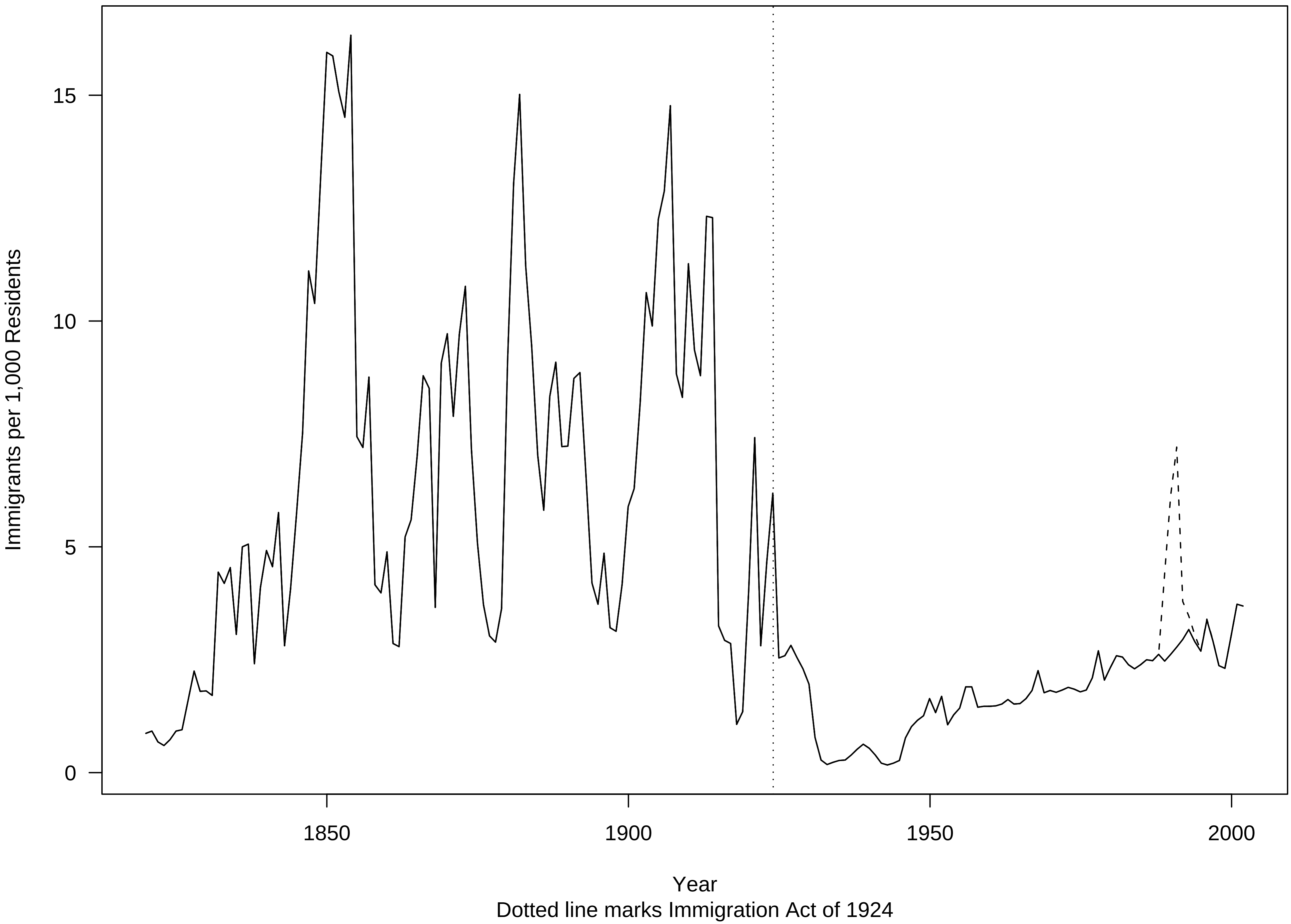


Figure 2. Legal Immigrants to U.S., 1820−2002
(Dashed spike indicates amnesty immigration)

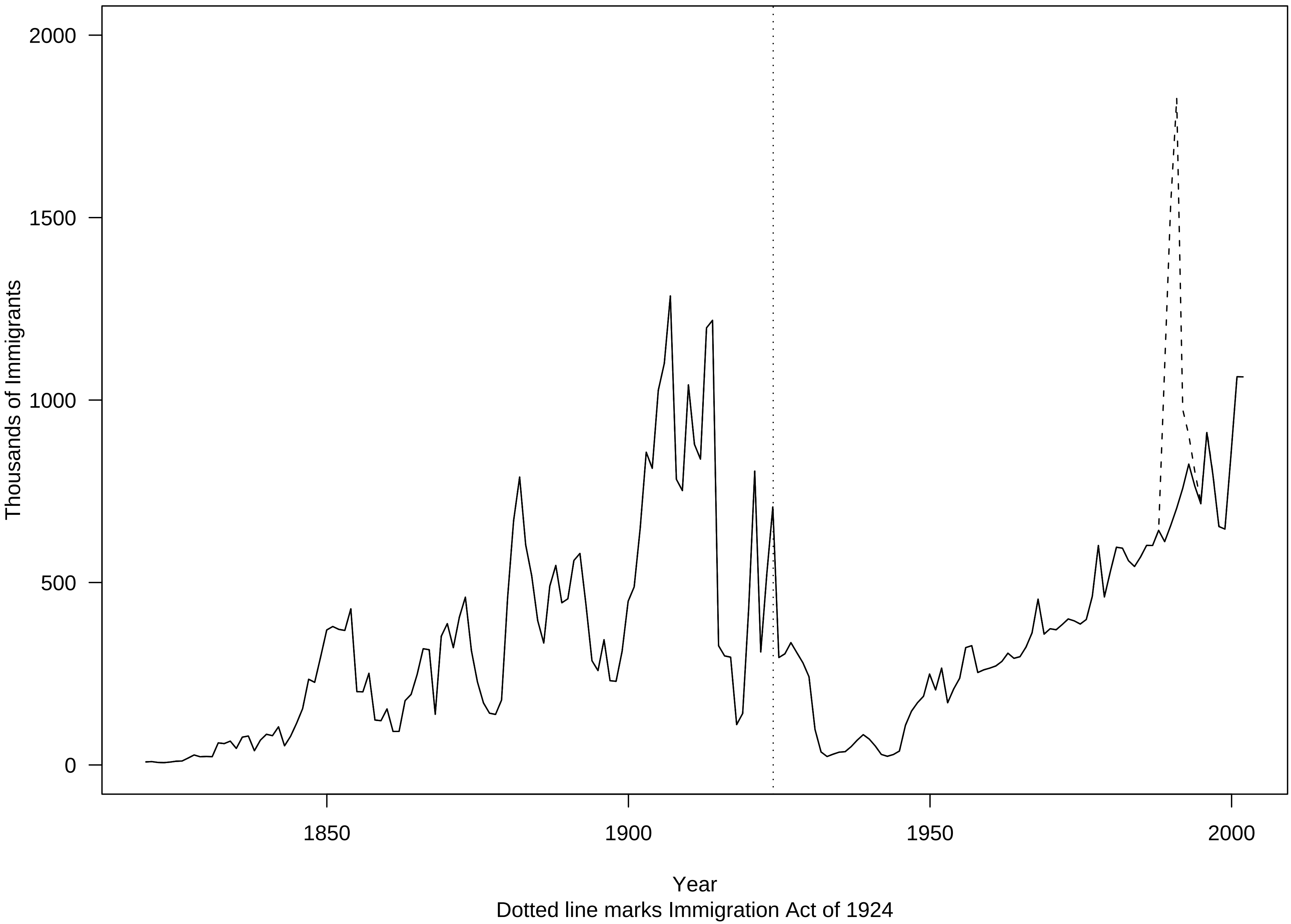


Figure 3. United Nations Estimate (1950-2000) and Forecast (2000-2050) of Net Immigration Rate to More Developed Region from Less Developed Region

Figure 4. Gross Legal Immigrants, 1820−2002 with Forecast of Net Immigrants
Stochastic Forecast without historical trend

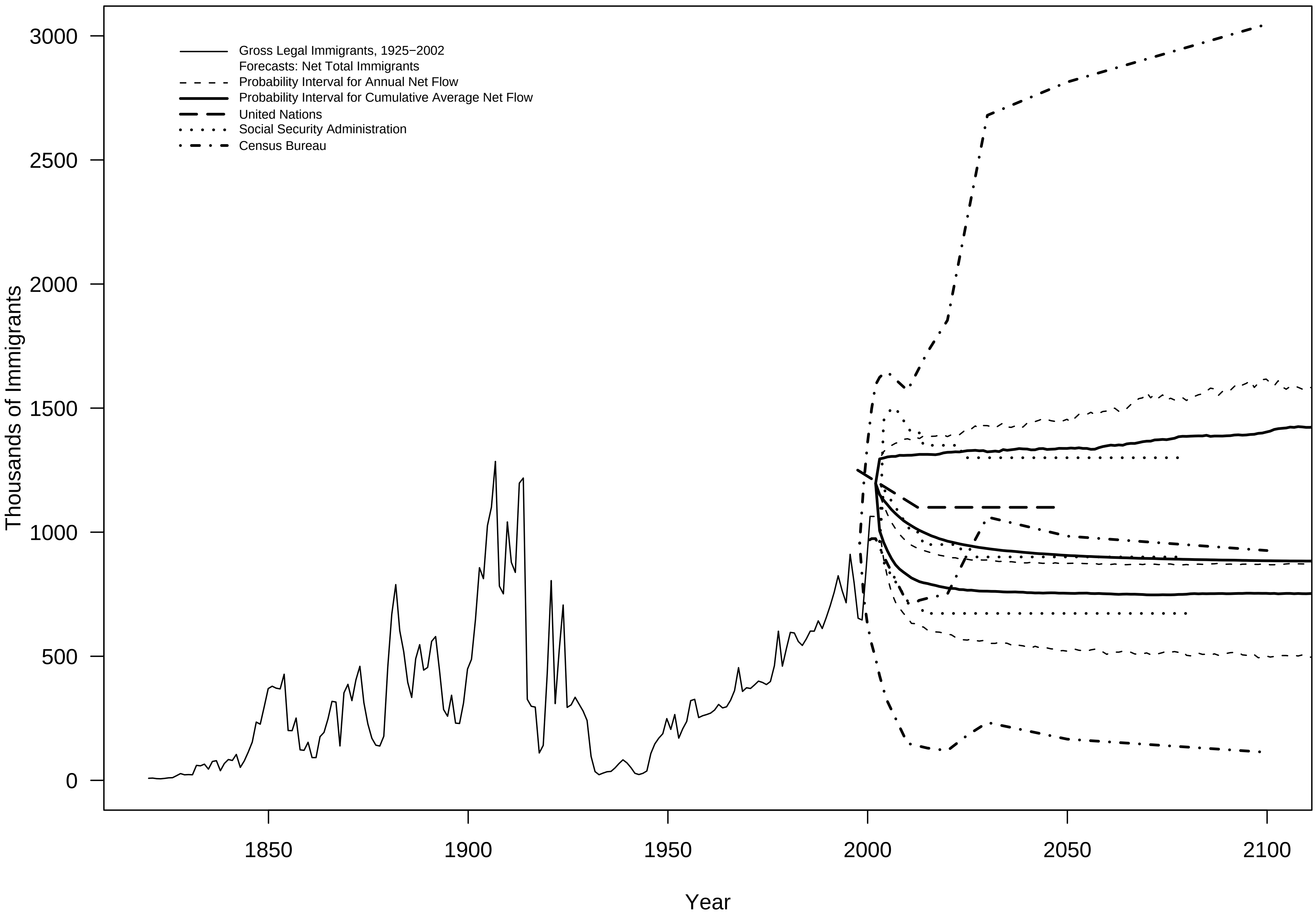


Figure 5. Gross Legal Immigrants, 1820−2002 with Forecast of Net Immigrants
Stochastic Forecast with historical trend

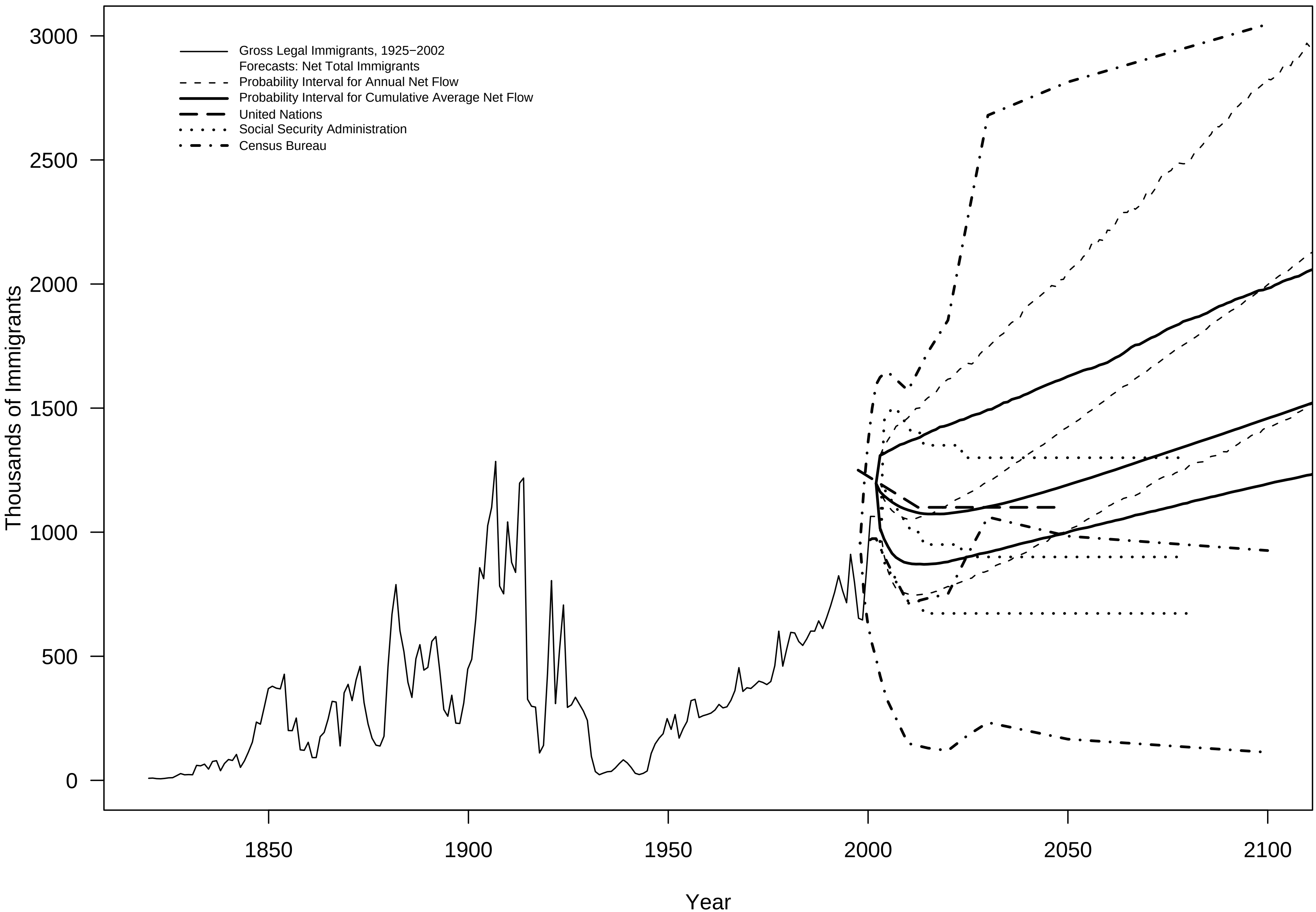


Figure 6. Gross Legal Immigration Rate, 1925−2002 with Forecasts of Net Immigration Rate
Stochastic Forecast without historical trend

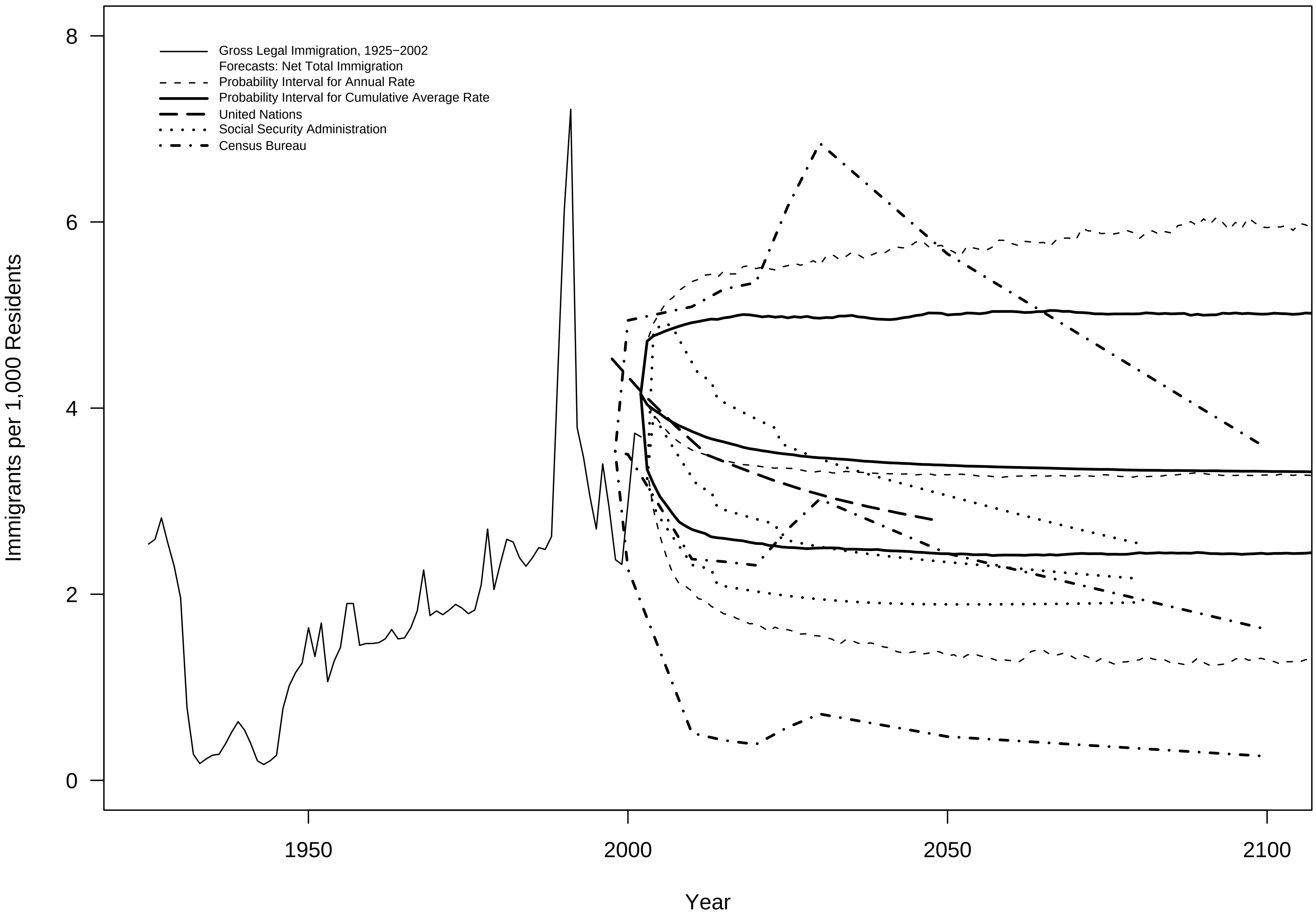


Figure 7. Gross Legal Immigration Rate, 1925−2002 with Forecasts of Net Immigration Rate
Stochastic Forecast with historical trend

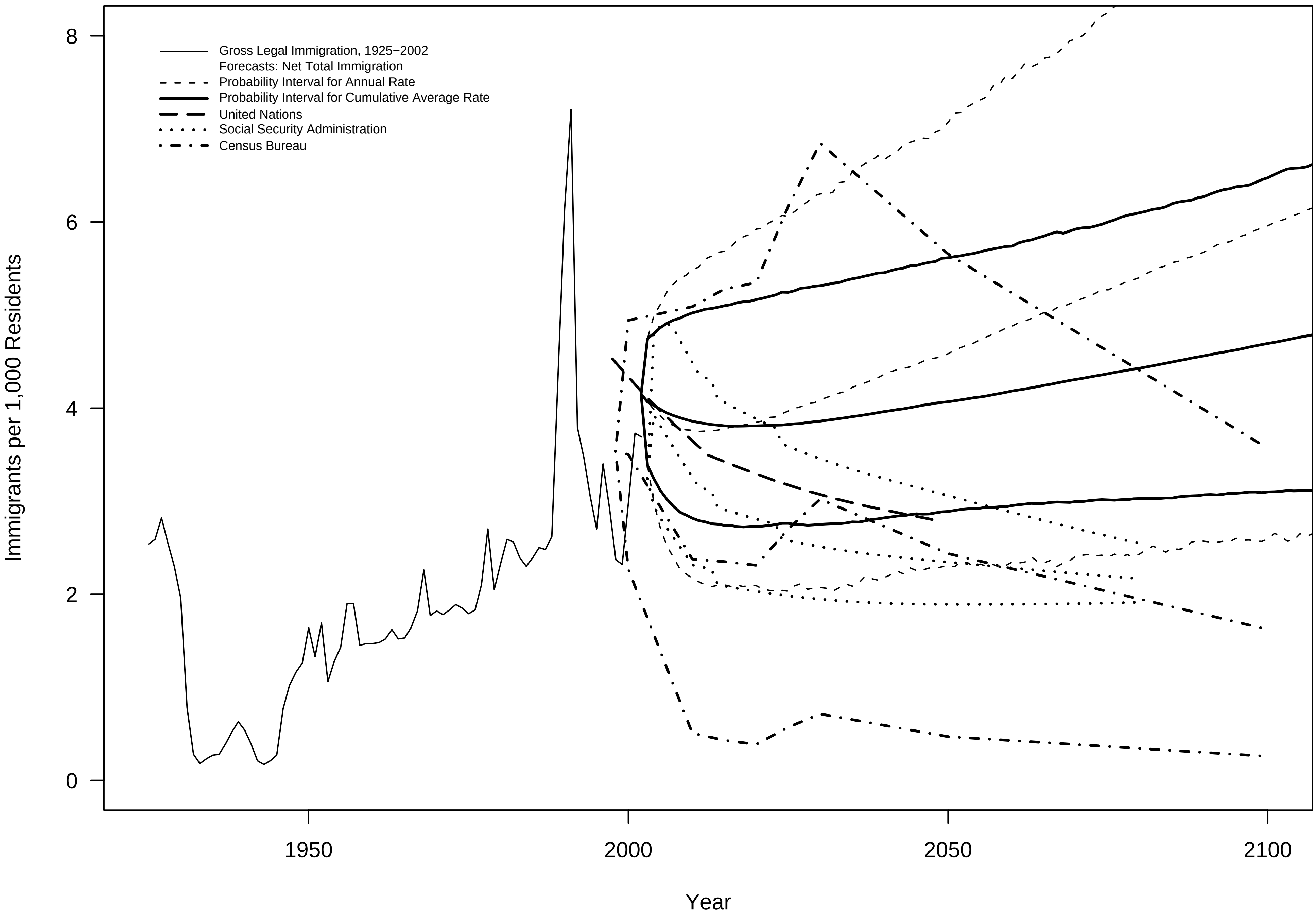


Figure 8. Population: median and 95% interval

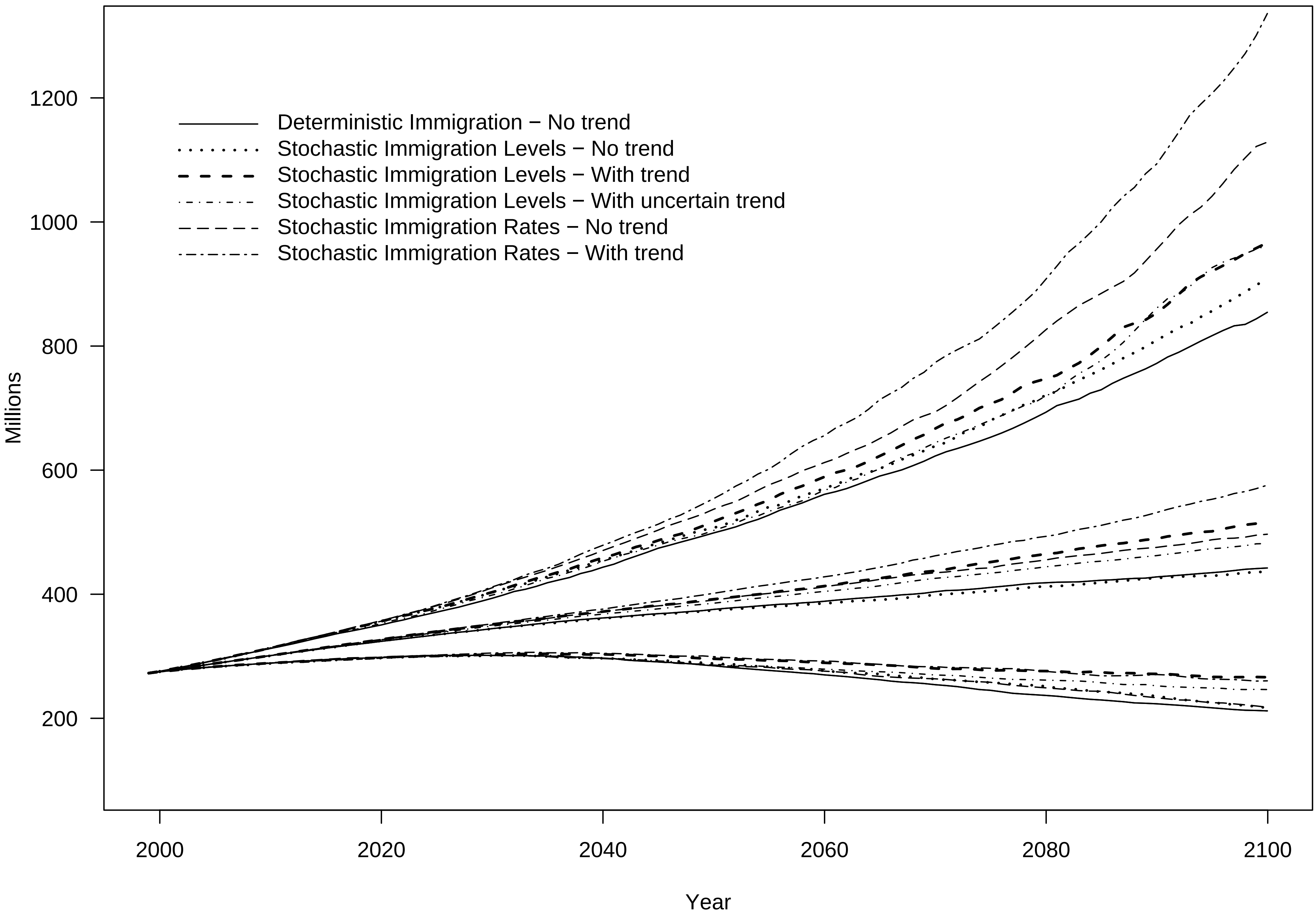
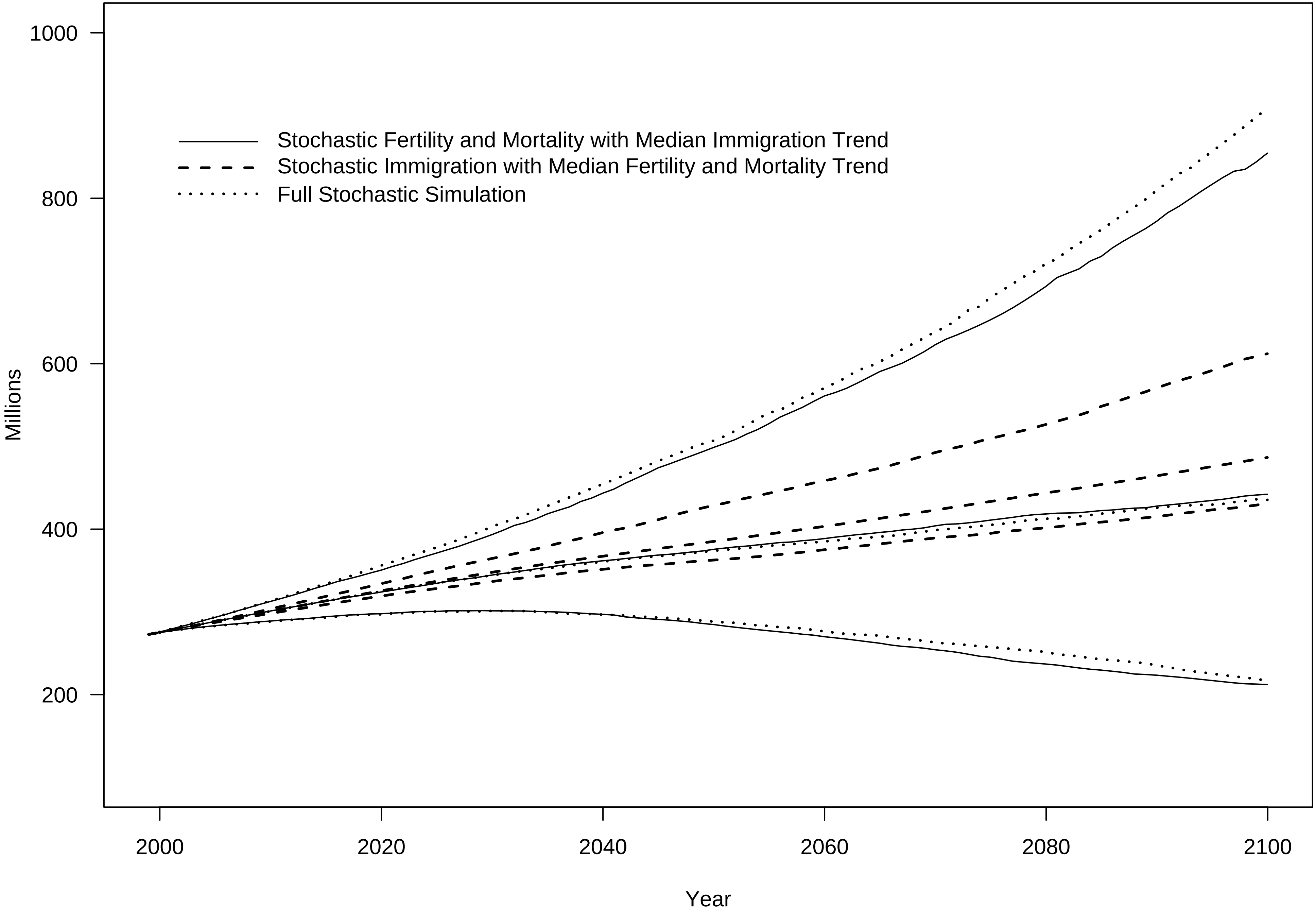


Figure 9. Effect of Stochastic Demographic Components on Population Forecast



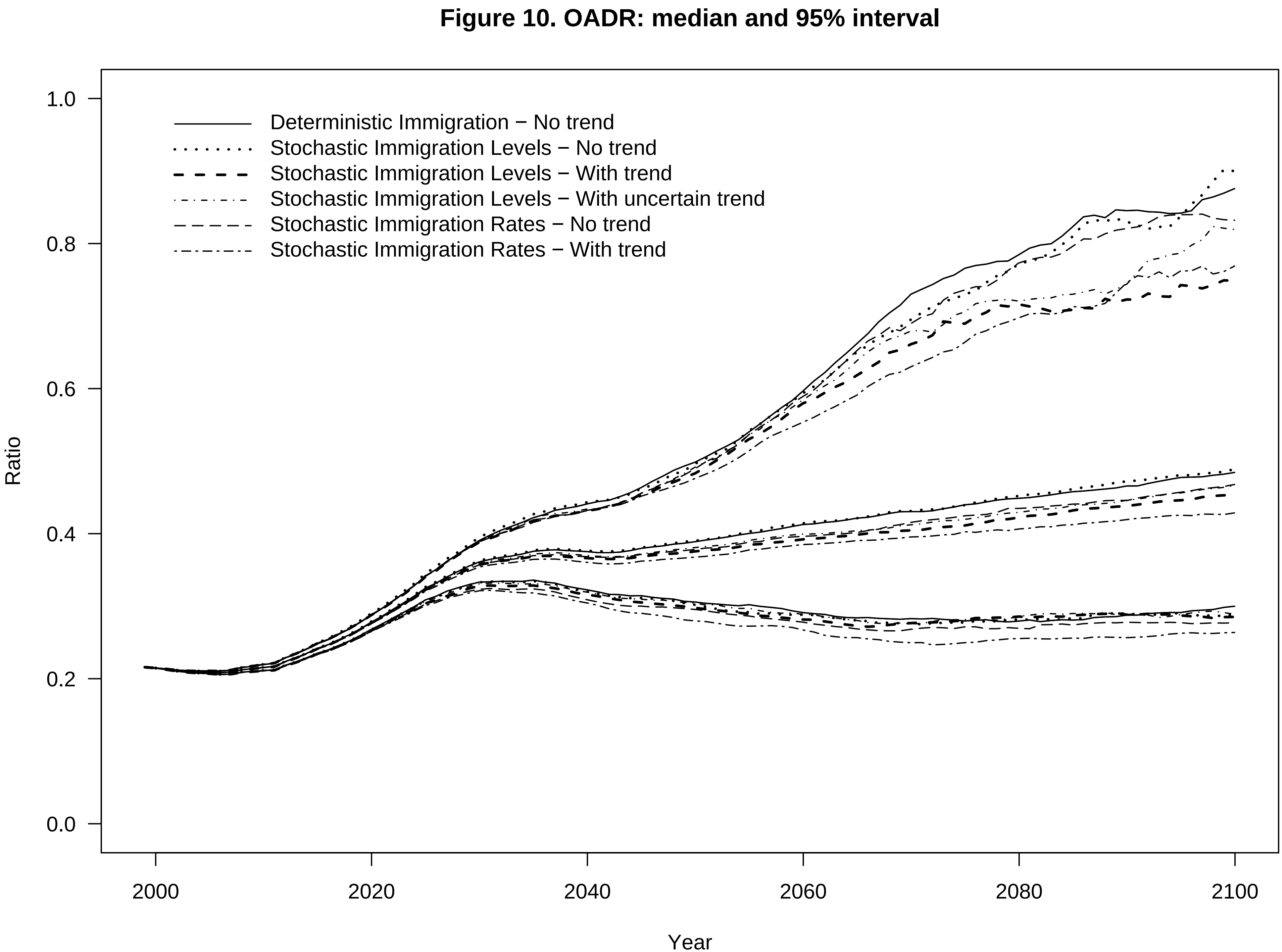


Figure 10. OADR: median and 95% interval
Deterministic Immigration − No trend
Stochastic Immigration Levels − No trend
Stochastic Immigration Levels − With trend
Stochastic Immigration Levels − With uncertain trend
Stochastic Immigration Rates − No trend
Stochastic Immigration Rates − With trend
Ratio
Year
0.0
0.2
0.4
0.6
0.8
1.0
2000
2020
2040
2060
2080
2100

Figure 11. Foreign born as a Percent of Total and Working−age Population

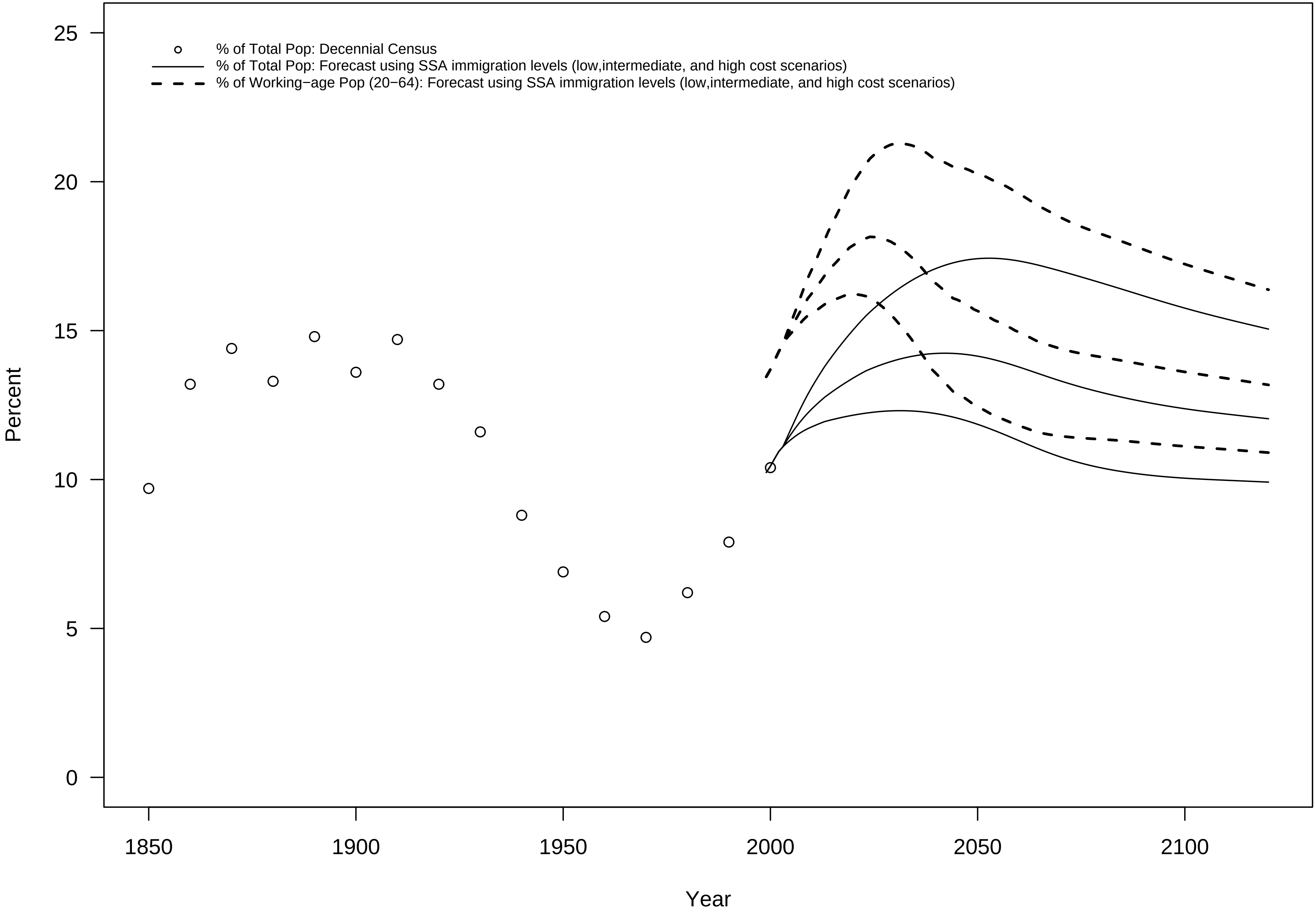


Figure 12. Gross Legal Immigrants, 1820−2002 with Forecast of Net Immigrants
Stochastic Forecast with different assumptions about the trend