ETLA
ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900
Telefax 358-9-601 753 World Wide Web: http://www.etla.fi/

Keskusteluaiheita – Discussion papers

No. 826

Juha M. Alho*

THE POPULATION

OF FINLAND IN 2050 AND BEYOND**

* Professor of Statistics, University of Joensuu, Department of Statistics, Box 111, 80101 Joensuu, Finland.

** This study was financially supported by the ageing programme of the Academy of Finland and by the Foundation for Municipal Development (Kunnallisalan kehittämissäätiö).

ISSN 0781-6847 11.11.2002

ALHO, Juha M., **THE POPULATION OF FINLAND IN 2050 AND BEYOND.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2002, 28 p. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; No. 826).

ABSTRACT: During the next fifty years, the growth of the population of Finland is expected to slow down, and turn into decline. The age-distribution is expected to become older because mortality declines. In particular, the share of the working age population will decline. This development is accentuated by the fact that the large post war birth cohorts reach retirement age during the next decade. Although there is general agreement about these broad features, it is difficult to say exactly when a decline might begin, or how high the age-dependency ratio will be like in the future. A study of past forecasts shows that demographic developments have repeatedly taken forecasters by surprise. We show that the forecast errors in Finland have not been related to other social phenomena, such as wars or economic crises, in a simple way. In fact, our understanding of the causes of the past errors is poor. Therefore, it is important to recognize that our current view of future population development may similarly be in error. We account for the uncertainty by statistical modeling. In this way it is possible to estimate, how large errors one should expect, if future demographic development is as volatile as in the past. Using stochastic simulation, we derive a predictive distribution for the future population vector that gives a realistic indication of the uncertainty to be expected.

KEYWORDS: Population forecasting; predictive distributions; demography.

ALHO, Juha M., **THE POPULATION OF FINLAND IN 2050 AND BEYOND.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2002, 28 s. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; No. 826).

TIIVISTELMÄ: Seuraavana viitenäkymmenenä vuotena Suomen väestön kasvun odotetaan hidastuvan ja kääntyvän laskuun. Ikärakenteen odotetaan vanhenevan, koska kuolevuus laskee. Erityisesti työikäisen väestön osuus vähenee. Tätä kehitystä voimistaa sodan jälkeen syntyneiden suurten ikäluokkien eläkkeelle siirtyminen seuraavan vuosikymmenen kuluessa. Vaikka odotetun väestönkehityksen pääpiirteistä onkin laaja yksimielisyys, on vaikea tarkasti arvioida, milloin väestön väheneminen alkaa tai kuinka korkeaksi huoltosuhde tulee muodostumaan. Aiempien Suomen ennusteiden tarkastelu osoittaa, että väestön kehitys on toistuvasti yllättänyt ennustajat. Osoitamme, että ennustevirheet eivät ole millään yksinkertaisella tavalla johtuneet esimerkiksi sellaisista muista yhteiskuntailmiöistä kuin sodat tai taloudelliset kriisit. Ymmärryksemme menneiden ennusteiden virheiden syistä onkin varsin vaatimatonta. Tästä syystä on tärkeää tunnustaa, että nykyiset käsityksemme tulevasta väestönkehityksestä voivat osoittautua virheellisiksi. Käytämme tilastollisia malleja menneen väestöepävarmuuden kuvaamiseen. Tällä tavoin on mahdollista arvioida, kuinka suureen epävarmuuteen nyt olisi syytä varautua, jos väestönkehityksen volatiliteetti on samalla tasolla kuin ennen. Käyttämällä stokastista simulointia johdamme prediktiivisen jakauman tulevan väestön muodostamalle vektorille. Jakauma antaa realistisen kuvan odotettavasta väestöepävarmuudesta.

ASIASANAT: Väestöennusteet; prediktiiviset jakaumat; demografia.

Contents

1. INTRODUCTION

During the next fifty years life expectancy in Finland will rise to an unprecedented level, but it is hard to say exactly, how high. Births are expected remain at a low level, but no-one knows whether fertility will remain constant, go down to the level now experienced in some Mediterranean countries, or whether it might rise to the level of the United States, or higher. Despite the uncertainty, it is likely that deaths eventually exceed births, and the population will turn into a decline, unless there is compensating migration into the country. Net migration is expected to be positive, but since it may be influenced by the EU policies, political upheavals in distant countries, or by surprises of the economy, few are bold enough to claim that they know what the level will be even a few years from now.

The age-distribution is becoming steadily older. In particular, the very oldest age-groups will experience unprecedented growth. Demand for all forms of elderly care will rise. On the other hand, those who are born in the near future may belong to small cohorts, have plentiful day care services, benefit from excellent teacher-to-student ratios in schools and universities, and eventually inherit comfortable dwellings and other forms of wealth. This is much more uncertain, however.

Despite their informativeness, qualitative characterizations, such as those given above, are dissatisfying in that for many planning purposes more precise descriptions would be preferable. Examples include pension systems, health care, social services, and education.

As we cannot know future events without error, we have to settle for less. Statistical modeling and techniques from time-series analysis provide us a language that can be used for the task (e.g., National Research Council 2000). The idea is a simple one. Although we cannot know what life expectancy is fifty years from now, we may express our uncertain knowledge regarding this value in terms of a *predictive probability distribution*. The median of the distribution is a value such that we consider it equally likely that the true future value will be above it and below it. The spread of the distribution around the median describes the magnitude of uncertainty. A full description of the possibilities can be given by providing the forecast user the entire predictive distribution. A more summary description can be given in terms of percentiles, deciles, or quartiles (cf., Törnqvist 1949).

Using probabilistic language has the disadvantage that forecast users must be familiar with the basic concepts of probability theory. As human intuition concerning probabilities is often weak, the presentation of probabilistic forecasts to non-experts poses a particular challenge. Graphical displays may be the most intuitive way to convey the information. On the other hand, a major advantage of probabilistic descriptions is that concepts from *risk analysis* and *stochastic control theory* can be brought to bear on social decision making. Issues of risk aversion and calculated risk taking become relevant parts of the political process.

In this forecast of Finland's future population we use probabilistic language systematically. All assumptions about the future will be made in terms of predictive distributions for the vital processes of fertility, mortality, and migration. A computer program PEP (Program for Error Propagation) is used to derive the consequences of the assumption for the population itself. PEP uses stochastic simulation to do that. In the current exercise, 3,000 simulation rounds were used. In other words, the whole population forecast was computed 3,000 times with randomly varying values for the vital rates of fertility, mortality, and migration. In recent years, empirically based stochastic forecasts have been made for the United States, e.g., by Alho and Spencer (1985), and Lee and Tuljapurkar (1994), and for Norway by Keilman, Pham and Hetland (2002).

As with any application of stochastic models, it is necessary that an understandable mapping between the events we observe and the model parameters be given. In the case of population forecasting this means that we must be able to defend both the center of each predictive distribution (or, why the median is where we put it?), and the amount of spread around the center (or, how much uncertainty should we expect?). Both aspects are non-trivial. In particular, while it is easy to acknowledge, in principle, that the future is uncertain, most people are surprised by the magnitude of uncertainty when it is presented to them.

To understand how the uncertainty arises requires a longer term perspective than is customary in demographic analyses underlying conventional forecasts. Therefore, before explaining the forecast assumptions, in Section 2 we discuss two examples of past population forecasts in Finland. The first goes back as far as our current forecast period goes forward, or for over 60 years. The second goes back half this time, or thirty years. We use those early contributions as backdrops to illustrate how unexpected the subsequent development of the society can be. Yet, the upheavals need not have any direct relationship with demographics.

In Section 3 we present details about the predictive distributions of the vital processes. In Section 4 we provide some summary results of the forecast for the years 2002-2050, and a more limited discussion of what to expect in 2050-2065. This is an updated version of the forecast described in Alho (1998).

2. FORECASTING IN A CHANGING SOCIETY

The statistical modeling of demographic events is based on the assumption that the events are realizations of stochastic processes. The particular paths that fertility, mortality, and migration take, may never repeat themselves, yet some aspects of the processes are assumed to recur. Our forecast period is 2002-2050, and we even consider some aspects as far as 2065. What might be a relevant data period, and how should we think of future uncertainty?

Many demographers believe that the social and economic factors underlying demographic events change fast. Therefore, data from, say, two or more decades ago may already be deemed irrelevant in forecasting. I agree with this assessment, in the sense that the most recent past is the most relevant, when we attempt to specify the median of the predictive distribution. However, a restricted data period cannot give us reliable information of even “business as usual” type of random variation (cf., National Research Council (2000), p. 195), let alone the kinds of structural breaks that occur in history (cf., Hoem 1973). They, in turn are a major source of uncertainty. For example, in the case of Finnish mortality, there is year-to-year variation that has stayed fairly similar for decades, but there are also occasional changes of trend (Alho 1998, pp. 19-21). A predictive distribution for the years 2002-2050 must acknowledge the possibility of both. Speaking in favor of a long data period is the fact that statistical model choice is on a firmer footing, if more data are available. In particular, one may be able to quantify modeling error (cf., Draper 1995). Estimation error can also be reduced, if all available data are used.

Empirical studies of past forecast errors (e.g., Keilman 1990, 1997) suggest that official forecasts of population are much more uncertain than is generally believed. In particular, the decline in mortality has been underestimated, and future fertility has been overestimated during the post World War II period. Evidence from the United States (Cheeseman-Day 1993) indicates that there may not be general improvement in forecasting accuracy in countries whose basic demographic data are accurate.

The quality of the Finnish population forecasts has not been systematically studied. The existing evidence (Hämäläinen 1987) suggests that the situation is similar here as elsewhere. To develop a sense of how errors in forecasting come about, and how they may be related to other social phenomena, we consider two examples in different historical settings.

2.1. Modeen's Forecast in 1934

In 1934, the population statistics of the year 1930, and mortality statistics of the 1920's had been completed (Luther 1993). Gunnar Modeen, then actuary at the Central Statistical Office, prepared a forecast for Finland, with 1930 as the starting or jump-off year (Modeen 1934a). The forecast extended to the year 1980.

At the time, the declaration of independence, and the civil war, were less than two decades behind. The great depression had just been experienced. Finland was an agricultural society in the process of building a national identity. A practical motive for making a forecast was the concern that population growth might stop.

In Finland, Modeen was the first to apply the cohort-component method pioneered by Cannan (1895) for England and Wales. In this method each age is separately survived using age-specific survival proportions, and births are generated using age-specific fertility rates. The method had become a standard tool in population forecasting in Europe, during the 1920's and 1930's (DeGans 1999).

Modeen's forecast showed that the population would peak at just under four million in the 1970's. The actual population was 4.8 million at that time, so the error incurred during a 35 year forecast period was about 17%. What went wrong?

Underlying Modeen's assumptions were comparisons to England, France, Italy, Denmark, Sweden, Norway, and Germany. He assumed that mortality would remain at the level estimated for the years 1921-1930. Assumption I for fertility postulated that the absolute number of births remains constant. Assumption II postulated that the general fertility rate would remain at the average level of the years 1931 and 1932. A well-known Swedish demographer Wicksell (1934) used similar assumptions in a forecast for Norway. The Statistisches Reichsamt (1930) had also done so in Germany. Following the practice in other countries, Modeen assumed future net migration to be zero.

Although Modeen considered his calculations as being somewhat hypothetical, the next Finnish national forecast prepared by Modeen and Fougstedt (1938) also assumed that mortality would remain at a constant 1931-1935 level. They thought that even if a further decline were likely ”it is very difficult to predict, how big this decrease is in different age groups”. The three assumptions concerning births were also similar to the earlier ones. Either the total number of births was assumed to remain constant; or fertility was assumed to remain at the 1931-1935 level; or fertility was assumed to decline by 1% per year, in each age. Again, zero net migration was assumed.

In retrospect, we know that during 1935-1970 Finland actually lost approximately 300,000 inhabitants due to migration. However, this was not the primary reason

Modeen's forecast was grossly in error. In fact, the zero net migration assumption tended to compensate for other errors!

Apart from the effect of the civil war in 1918, a stable declining trend in mortality starting around 1870 had been observable. Despite, neither Modeen's forecast nor that of Modeen and Fougstedt considered the possibility that the trend might continue. War casualties during World War II offset the resulting error to some extent, but still the forecasted mortality was much higher than the observed mortality. During 1935-1970 the net effect of the error was roughly 600,000.

When writing about fertility, Modeen (1934b, pp. 361-362) mentioned a French forecast in which a decline to the level of the department of Seine, ”a fairly urbanized area”, was assumed. Modeen dismissed this and other such hypotheses as being based more on ”speculation” than his own assumptions of constant fertility. Later, Modeen and Fougstedt (1938, p. 8) argued that the urbanization of the society will most likely lead to a decrease both in marriage rates and in marital fertility. This was the prevailing view elsewhere, as well. In Germany, Burgdörfer (1932, p. 32) wrote of Berlin as the ”unfruchtbare Stadt” and worried about the unhappy consequences of the ”Zweikinder-system”. In Sweden Myrdal and Myrdal (1934, pp. 87-88, 94) attributed the decline in fertility to improved contraception, and secularized rationality that follow from urban life style. They thought the decline would continue for the ”nearest decades”. Later, Whelpton (1947, pp. 28-29) argued similarly that the U.S. fertility would continue to decline, because of urbanization, women's increased labor force participation, and improved contraception.

As we know now, all these authors turned out be in error. In the case of Modeen, the net error during 1935-1970 was approximately 300,000 too few births.

Since the three error components are not independent, no unique decomposition of the total error exists. Nevertheless, ignoring the interactions, one can say that the three error components, as identified here, explain -300,000 + 600,000 + 300,000 = 600,000 inhabitants out of the total of 800,000 by which the true population in 1970 exceeded Modeen's forecast.

For later analysis, three aspects are notable. (1) The level of error in all three components was extremely high. Hence, the possibility of errors should be taken seriously in forecasting. (2) Error components compensated for each other, to a considerable extent. Probabilistic methods are capable of handling the independence (or lack of independence) of error sources. (3) Mortality was drastically overestimated. Overestimates have continued, but their magnitude has been lesser as the later forecasts have taken declining trends partially into account.

2.2. Population Plans of the 1970's[1]

After World War II, Finland's economic, political, and social life was quite different from the time Modeen made his pioneering forecast. After the heavy war reparations, the 1960's were a period of rapid growth with an (inflation adjusted) increase of 55% in Gross Domestic Product per capita during 1960-1970 (Hjerppe 1989). At the same time manufacturing became the dominant sector of the economy, and the country urbanized rapidly. Many thought that “rational planning” would solve the problems of the society. Despite the ideological differences, a major concern in the 1970's, just like in Modeen's

---

[1] The author would like to thank Professor Heikki Eskelinen for pointing out some of the central references to this period. Any interpretations given to the texts are solely the responsibility of the author.

time, was the slowing down of population growth. The total fertility rate had declined from 2.7 children per woman in 1960 to 1.8 in 1970. Emigration exceeded immigration, and Finland suffered a net loss of 187,000 inhabitants during the years 1960-1970 due to this. Neither course of events could be forecasted even as late as the early 1960's (Väestöennusteryhmä 1973, p. 21).

In accordance with the general mood, the government decided to replace "by-stander's forecasts" of population that incorporate no assumptions about specific policies, by "participant's forecasts" in which the state would harmonize social policies on the regional level in such a way that future population would actually follow a *population plan* (Väestöennusteryhmä 1973). Conceptual models for the work were sought from the regional input-output tables and other planning tools developed in Sweden, Norway, Italy, France, the Netherlands, the United Kingdom, West-Germany, and the Soviet Union. These models attempted to give a system theoretic picture of the regional economies, regional populations, and their change (Talousneuvoston aluejaosto 1972). The primary interest of the planners was in the interregional migration. Yet, it is of interest to see how the general demographic patterns were viewed at the time.

Forecasts made in the late 1960's by Statistics Finland (then the Central Statistical Office), and by the regional planning authorities, put Finland's population in 2000 between 5.15 and 5.65 million (Väestöennusteryhmä 1973, p. 19). A forecast by Statistics Finland was the lowest, and hit almost exactly the observed value of 5.18 million. However, in view of the events of the late 1960's, in the early 1970's it was felt that these forecasts were too optimistic. In 1971 Statistics Finland produced a new forecast reflecting the recent lower fertility and higher out-migration. The change of mood was remarkable. The new forecast was 4.4 million for the year 2000. This meant that there was an error of 15% in thirty years.

In 1974, the forecast was revised upwards again. Four population figures for the year 2000 were given, between 4.5 and 4.9 million (Population Projections 1973-2000). In three years, pessimism regarding migration had decreased and zero net-migration was assumed. Mortality was either assumed to stay at the level of 1973, or to decline slowly until Sweden's levels of 1971 would be reached. In the case of fertility there was more variation. Either it was assumed to remain at the 1972 or 1973 level; or to decline by about 20% in a decade and then stabilize; or to remain first at the 1972 level, and then increase to replacement level by the year 2000. Clearly, the forecast methodology had improved from that used a few years earlier. Yet, the median of the four forecasts for the year 2000 was 4.6 million, an error of still 10% in 27 years. The prevailing rationalistic outlook was of little help in forecasting.

For the 1974 forecast, the cumulative net migration forecast was 80,000 too low by the year 2000. The number of deaths was approximately 200,000 too high, and the number births was (for the median of the forecasts) approximately 230,000 too low. Ignoring interactions of the three components, this produces an error of approximately 80,000 + 200,000 + 230,000 = 510,000 that covers most of the actual error of 580,000 for the year 2000. As in the case of Modeen, all errors were large, and mortality was overestimated. However, this time there was no cancellation of error.

2.3. Uncertainty in the 21st Century

After the success of Nokia and other information technology companies, many may feel that the factors that lead Modeen astray, or the factors that lead the planning optimists of the 1970's astray, would not be relevant at the turn of the millennium. After all, instead of nationalism, globalization is the topic of debate. Instead of planning, great trust is put into market mechanisms. Finland has joined the European Union, and opened up its economy. Curiously, the prospect of a declining (and hence, aging) population is again a concern. How should we think about the uncertainty of the future now?

Between Modeen and the plans of the 1970's, World War II took place. After the 1970's the Soviet regime fell out of power. Momentous as these events were, the errors of the population forecasts cannot be attributed to them. To be sure, World War II lead to an increase in mortality, but Modeen's forecast of mortality was still too high, so the war actually made the mortality forecast more accurate. The role of the war in the creation of the baby-boom is also debatable, because, contrary to what many people believe, it cannot be explained simply as a matter of recovering births that would have been postponed during the war. Since this is a matter of more general concern, we present, in an Appendix, a simple "back-of-the-envelope" calculation that proves the point.

Moreover, although the demise of the Soviet Union did open up the borders of Russia to out-migration, this did not lead to a mass migration to Finland that one might have expected.

The connection between the political and economic upheavals Finland has experienced in recent decades, and the demography of the country, appears rather weak. Therefore, instead of speculating about the kinds of political surprises we might encounter in the future, it seems preferable to use past demographic fluctuations directly as a guide to the fluctuations to be expected in the future. It has been demonstrated earlier (Alho 2000) that in the 1930's, when Modeen made his forecast, a stochastic analysis of past data could have been used to produce a predictive distribution that would have correctly indicated the level of error one should have expected at the time. (Of course, neither the appropriate theory nor the computational facilities existed at the time.)

3. FORECAST ASSUMPTIONS

We will now sketch the assumptions concerning the predictive distributions that are needed in a cohort-component setting.

In addition to forecasting to 2050, we had an interest in longer term calculations that extended until 2065. The quality of the existing error estimates begins to deteriorate when the forecast period exceeds 40-50 years. In this forecast, the level of uncertainty in the forecasting of fertility and mortality was assumed not to increase after 50 years (cf., Alho and Spencer 1997, p. 218).

3.1. Jump-off Population

The starting, or the jump-off population was the Finnish population at the end of year 2001, by single years of age (0,1,2,..., 119, 120+), for females and males. The population being forecasted is the legally resident population as enumerated in the Finnish population register. The total count was 5.195 million.

The population register is thought to be nearly 100% accurate due to the administrative uses of the population registration system. Complete accuracy was assumed in this forecast. In countries with less accurate registration systems, the uncertainty of the jump-off value can also be expressed in probabilistic terms. For an example concerning the United States, see Alho and Spencer (1985).

3.2. Fertility

The median of the distribution of fertility was assumed to be at the level observed empirically in 2001, in each child-bearing age (14-50, in this application). A summary measure used to describe the overall level of fertility is the total fertility rate. It is defined as the sum of age-specific fertility rates over the child-bearing ages. The total fertility rate of a given year can be interpreted as the expected number of children a woman will have provided that (a) she experiences the age-specific rates of the year in a sequence, and (b) she survives to the end of the child-bearing ages. In this case the total fertility rate was assumed to remain indefinitely at 1.73, the level observed in the year 2000.

Comparing our forecast to the year 2000 revision of the U.N. forecast, the difference is that the U.N. (medium variant) first assumes the Finnish fertility to decline to about 1.55 in 2000-2005 and then to increase to 1.94 in 2045-2050 (United Nations, 2001, p. 589). This is in accordance with the U.N.'s general view that currently a decline of fertility is under way in Europe, but later all countries move towards a value guaranteeing population replacement (approximately 2.07). As there is little solid research to support such a development, we have opted for the simpler assumption. Numerically, and in practical terms, our assumption and that of the U.N. are quite close, however.

The uncertainty estimates used to create the predictive distribution of the future fertility and mortality in Finland are similar to those described in Alho (1998). They are based on a statistical analysis of the Finnish total fertility rate during 1776-1996. We have empirically determined the relative error of a naive forecast that assumes fertility to remain constant in the future. A naive forecast approximates closely the medium forecasts made in Finland (starting from Modeen's and continuing to the present day), and in the United States (e.g., Lee 1974). To assess the error of the naive forecasts, we used the medians of the absolute errors to determine the scale parameters. This means that the largest errors observed in the past did not influence the estimate. Furthermore, the estimates were scaled to match the recent low level of error in the naive forecasts. The estimates were also compared to error estimates obtained from formal time-series (ARIMA) models. The latter indicate a larger level of uncertainty, but this is due to their constant volatility assumption and lack of robustness against outliers, so the lower estimates are preferred. The estimates were also checked against those obtained from the period 1967-1996 when fertility was relatively stable. The shorter period could not serve as a basis for error estimation in long term forecasting, but it turned out that for short forecast periods the two approaches produced similar estimates of uncertainty (Alho 1998, p. 18).

In summary, the proposed error structure corresponds to recent past uncertainty during 10-20 years into the future. After that it corresponds to the historical median levels of uncertainty from 1776-1996.

The predictive distribution is graphically displayed in Figure 1. The median (Md), the first and third quartiles (Q_1, Q_3), and the first and ninth deciles (d_1, d_9), for the years 2030 and 2050 are as follows:

year	d_1	Q_1	Md	Q_3	d_9
2030	1.21	1.43	1.73	2.08	2.46
2050	1.10	1.36	1.73	2.20	2.72

The intervals of Figure 1 (and those of all subsequent figures) are annual. Joint intervals that would cover the whole path of total fertility would be wider.

3.3. Mortality

Mortality rates of a given year can be summarized in terms of life expectancy at birth. This is the length of life a person is expected to have, if he or she would experience the year's age-specific mortality risks in each age. In the year 2000 life expectancy for females was 81.0 and for males 74.1 years. To put these figures into a perspective, we note that in the early 1970's the life expectancies were about 75 and 67 years, and at the time of Modeen, about 57 and 52 years, respectively.

The point forecast for mortality was based on a simple trend extrapolation of the recent past age-specific mortality (in the log-scale) by single years of age, for males and females separately. To reduce variation caused by small numbers of events, the rates of change were smoothed by the procedure RSMOOTH of Minitab. This lead to the forecast of 86.7 years for females in 2050, and for 81.6 for males. Or, the gap between males and females is expected diminish. The U.N. forecasts 86.1 and 79.8 for females and males, respectively, in 2045-2050 (United Nations (2001), p. 633). Given the length of the forecast period, the difference between the U.N. forecast and ours must be considered small.

The analysis of uncertainty was based on the relative error of the naive forecast with data for 5-year age-groups for 1900-1994. In the case of mortality, the naive forecast assumed that the recent past decline in mortality continues indefinitely. As discussed in Alho (1990) and Lee and Miller (2001), the naive method and related extrapolation methods perform typically equally well as (or better than!) the judgmental official forecasts. The analysis of uncertainty is complicated by the fact that the level of mortality varies considerably by age. In ages with low level of mortality, very large relative errors are commonplace. The proposed estimates reflect the uncertainty in ages 35-59, 60-79, and 80+ correctly, but may underestimate the relative error in the younger ages. This is not a great concern, however, since the absolute numbers of deaths are small in those ages. The reasonableness of the assessment of uncertainty of mortality was checked by a comparison to a formal time series model. In this case the agreement was notably good (Alho 1998, p. 22).

The implications of these estimates have been summarized in terms of the predictive distributions for life expectancy in Figure 2. Summary data for the years 2030 and 2050 are as follows:

sex	year	d_1	Q_1	Md	Q_3	d_9
female	2030	82.7	83.8	85.0	86.2	87.3
	2050	83.3	84.6	86.7	88.4	90.1
male	2030	75.9	77.5	79.3	81.0	82.6
	2050	76.7	79.0	81.8	84.3	86.4

Lee and Carter (1992) used a time-series approach to forecast the U.S. life expectancy (both sexes combined). The resulting prediction interval for a 50 year ahead forecast had width of about 8.4 years. For comparison, a 95% prediction interval for the Finnish females is [82.4, 91.1] and for males [75.3, 87.9]. Or, the widths are 8.7 and 12.6, respectively. Keilman, Pham and Hetland (2001, pp. 48-49) report empirical analyses with Norwegian data. Their 95% prediction interval 50 years ahead has width of about 9 years for females and 12 years for males. A possible reason for why one would expect greater uncertainty in the Nordic countries as compared to the United States, is that the U.S. consists of several large, culturally somewhat independent sub-populations, so on average, one would expect cancellations of fluctuations, and thus, a more stable development.

Recently, Vaupel and Oeppen (2001) have presented evidence of the development of the so-called “best practice life expectancy”. This is the life expectancy of the country that at any given time has the longest life expectancy. Vaupel and Oeppen show that for females the curve goes almost linearly from the value of 45 years observed in Sweden in 1840, to 85 years observed in Japan in 2000. Or, life expectancy has improved by approximately 0.25 years annually. In 2000, the Finnish female life expectancy was 81.0 years, a difference of four years to the Japanese. If Finland would experience improvements at the same rate as the best practice country, in 50 years we would expect female life expectancy to be 50/4 = 12.5 years higher than now, or 93.5 years. This value is higher than the 9th decile of our predictive distribution. How should we think about the discrepancy?

From Modeen’s time, or in about 70 years, female life expectancy increased by 81 - 57 = 24 years. This is approximately 0.34 years annually, or more than the improvement in the best practice countries. However, when we look at the first 40 years of the period, the increase is 75 - 57 = 18 years, or 0.45 years, each year. During the latter 30 years the increase was 81 - 75 = 6 years, or 0.2 years, each year. The latter rate would imply an increase of 10 years by 2050, a value that is still above our 9th decile. Our forecast assumed that age-specific mortality rates continue to decline at the rate they have declined during the past 15 years. One can see from the Figures 5a-e of Alho (1998, pp. 19-21) that especially in the ages 80+ mortality has stagnated since 1980, or so. In fact, during the past 15 years female life expectancy improved by only 81.0 - 78.6 = 2.4 years, or 0.16 years annually. This would imply an improvement to 88 years by 2050. This is at 0.70 fractile of the predictive distribution. The difference between 88 years and the median of 86.7 can be fully reconciled by noting that the continuation of the recent rate of decline in age-specific mortality implies a slowing down of increase in life expectancy (Alho 2002). This, of course, is incompatible with the linear change hypothesis.

3.4. Migration

The forecasting of migration differs from that of fertility or mortality in at least three ways. First, migration can be influenced by government policies to a higher extent than fertility or mortality. Second, although out-migration can be reasonably analyzed via out-migration rates, it is typically difficult to define a meaningful risk population for in-migration. Third, data on migration are poor even in a country like Finland that has a well functioning population register. Because of these problems, migration forecasts are typically judgmental, and given in terms of the net number of migrants one expects. On the other hand, a probabilistic approach is well suited to the handling of the uncertainty of judgment concerning future migration. The primary difficulty is in finding a robust way to elicit judgments.

In Alho (1998) it was assumed that the most likely net number of migrants would remain at the recent level of 4,000 per year. A prediction interval around that number was derived using a combination of empirical data analysis and a betting argument. The argument was used to calibrate the empirical estimates to match the subjective probabilities of the experts of Statistics Finland (Alho 1998). Less emphasis was put on the long term implications of the assumptions. Here, we describe the assumptions from that perspective.

An annual gain of 4,000 during a 50 year forecast period would mean that the most likely value for the net gain is 50 H 4,000 = 200,000 inhabitants. In the current revision it was felt that the magnitude of gain may be too optimistic. Instead, it was assumed that net migration would start from the recent past value of 5,000, and decline to zero in 25 years. This implies 70,000 as the most likely gain. In Alho (1998) the autocorrelation structure of net-migration was chosen so that it would be, at all lags, equal to the first empirically observed autocorrelation, or 0.716. This is a conservative assumption. It implies that the standard deviation of the cumulative net-migration during 2001-2050 would be as high as 474,200. In this exercise it was felt that the value is somewhat too high. After some experimentation, an autocorrelation of 0.4 was chosen, a value closer to the average autocorrelation. This means that the assumed predictive distribution of the cumulative net-migration in the next 50 years would be $N(70{,}000, 409{,}800^2)$. Or, a 50% prediction interval for the cumulative net migration is [-206,400; 346,400], an 80% interval is [-455,200; 595,200], and a 95% interval is [-733,200; 873,200]. In contemplating whether this is a realistic assessment of the uncertainty, we remind the reader that Modeen's migration assumption was off by 300,000 in 35 years, and Statistics Finland's 1974 assumption was off by 80,000 in 25 years.

4. PREDICTIVE DISTRIBUTIONS

4.1. Total Population

Figure 3 has summary data of the predictive distribution of the year-end population in 2002-2050. These (and subsequent) results are based on 3,000 simulations of the population. The median may be taken as a point forecast. The first and third quartiles form the 50% prediction intervals, and the first and the ninth deciles form the 80% prediction intervals. The uncertainty of population growth vividly illustrated by the width of the intervals. According to the median, the population is expected to grow for less than a decade, to just above 5.3 million, and then a gradual decline begins. However, the median forecast is quite flat for over two decades. Relatively small changes in migration can make a difference, from this perspective. In fact, it is by no means certain that the population will turn into a decline, if net migration is positive, or fertility happens to increase.

Figure 4 gives a more detailed look at the predictive distributions in 2030 and 2050.[2] The histograms show how the uncertainty increases as we go further into the future. Summary statistics for the total population in these years are as follows (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	4,836	5,020	5,243	5,480	5,676
2050	3,947	4,361	4,867	5,469	6,022

Given the overall level of uncertainty small modifications in assumptions do not matter much. However, it is of some interest to observe that had we retained the earlier net migration assumption of 4,000 per year, it would have had the effect of increasing the median in 2030 approximately to 5,273,000, and in 2050 to 4,997,000. In other words, the expected decline would have been somewhat delayed as compared to the above.

An assessment of the roles of fertility and mortality on one hand, and of migration on the other, in the overall uncertainty, can be made by repeating the calculations, but setting the uncertainty of the migration forecast to zero. The following results are obtained. The standard deviation of the predictive distribution in 2050, when all sources of uncertainty are present, is 0.828 million. If there would be no uncertainty due to migration, the standard deviation would be 0.640. On the other hand, if only the uncertainty in the cumulative net number of migrants would be taken into account, the standard deviation would be 0.410. One can deduce from these estimates that, approximately, 60% of the variance is explained by fertility and mortality together, 25% is explained by migration, and 15% is due to nonlinearities and interactions.

In National Science Foundation (2000) a model was built for the errors of the U.N. national forecasts of all countries of the world, with data from the 1950's onwards. The ratio of the 9th decile to the median of Finland's predictive distribution in 2050 is 1.173. From the current calculation we get 6022/4867 = 1.237. It is somewhat reassuring that the two completely independent methods give similar estimates of uncertainty.

4.2. Age-distribution

Figure 6 presents a forecast of the population by age and sex in 2030, together with 80% prediction intervals. Figure 7 presents the corresponding information for 2050. Population aging is clearly visible in both graphs. Similarly, the uncertainty of fertility leads to wide prediction intervals in the younger ages. In adult ages, migration is a primary source of uncertainty. In the very highest ages the uncertainty of mortality is also an issue.

It is customary to summarize age-distributions in terms of dependency ratios. In Figure 7 we present the ratio of the population in ages 0-18 and 65+ to the population in age 19-64. Because of variations in labor force participation and unemployment this dependency ratio is not a pure measure of the burden posed by the economically inactive population on the economically active. Nevertheless, the ratio provides us with useful indication of the pressure the demographic development alone puts on the economy. We find that the burden will increase from the present value of 0.61 to about 0.88 in three decades. That is the burden increases by nearly a half. After that - when the baby boom

[2] The bin width was 200,000, and the percentages given on vertical axis refer to the number of simulation runs falling into a bin.

generations die - the burden is expected to ease out. Of course, this is just the most likely development. The summary statistics for age-dependency ratio in the years 2030 and 2050 are as follows:

year	d_1	Q_1	Md	Q_3	d_9
2030	0.78	0.82	0.85	0.90	0.94
2050	0.77	0.82	0.89	0.96	1.03

Looking at the ninth decile we find that the probability is 10% that the burden will be over 1.00 in about four decades. Moreover, even the first decile will increase to 0.78 in three decades. This means that even in very optimistic circumstances the future burden will increase at least by a third from the current value.

Because of the approaching retirement of the baby-boom generations, the role of the elderly in age-dependency is of interest. Figure 8 has a plot of the old age-dependency ratio, or the ratio of the population in ages 65+ to the population in ages 19-64. We see a dramatic increase from the present value of 0.25. The summary statistics for the old age-dependency ratio in the years 2030 and 2050 are as follows:

year	d_1	Q_1	Md	Q_3	d_9
2030	0.44	0.46	0.49	0.52	0.55
2050	0.41	0.46	0.52	0.58	0.65

4.3. Children and the Young

Current fertility is well below replacement level. We expect the low level to persist, but recognize that fertility has been highly volatile in the past. It might take us by surprise again. What do these facts imply for the births? Figure 9 has a predictive distribution for the births (actually zero-year olds at the end of the calender year). The median of the predictive distribution declines roughly linear from the current value of 55,000 to 47,000 by 2050. The summary statistics of the predictive distribution of future births in 2030 and 2050 are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	36.6	42.8	50.3	60.1	69.8
2050	23.9	33.2	46.9	66.5	91.3

We see that the probability is 10% that the births in 2050 are only half of the most likely value. On the other hand, with equal probability the births might be double the most likely value.

Consider the population in ages 0-6, the under school age population. Their absolute numbers are of interest in the planning of day-care services, for example. As one would expect, their forecast is almost identical to that of the births, albeit with a scale adjustment, see Figure 10. The summary statistics of the predictive distribution of the population in ages 0-6, in 2030 and 2050, are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	275	315	364	423	482
2050	177	241	331	462	618

The school-age population 7-18 is of interest to the planners of the educational system. Figure 12 presents the predictive distribution. The difference here is that initially the uncertainty is much less than in the younger ages, because for over a decade many of the members of the population subgroup are already born at jump-off time. We see from the graph that, demographically, pressures on the school system are likely to ease. The summary statistics of the predictive distribution of the population in ages 7-18, in 2030 and 2050, are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	534	593	666	747	823
2050	359	458	578	752	927

4.4. Working Age Population

Primarily due to the baby-boom generations, Finland has, in recent years, had a relatively large population in working age, which we define here as ages 19-64. Figure 12 has a predictive distribution for the size of that population. We see that the situation is going to get worse. Currently, the population size is 3,225,000. The summary statistics of the predictive distribution of the population in ages 7-18, in 2030 and 2050, are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	2,619	2,713	2,822	2,923	3,021
2050	2,101	2,320	2,560	2,858	3,133

4.5. Elderly and Oldest-Old

The retirement of the baby-boom generations has lead to much public discussion about the sustainability of the public sector finances, including pension systems. At jump-off time the size of the population in ages 65+ was 787,000. Figure 13 shows that the population is expected to nearly double during the next thirty years, or so. After that the growth is expected to stop, because the smaller cohorts born after the 1960's enter the age bracket. The summary statistics of the predictive distribution of the population in ages 65+, in 2030 and 2050, are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	1,259	1,318	1,386	1,449	1,507
2050	1,116	1,212	1,334	1,461	1,574

The oldest-old have become a topic of much research in recent years. In many countries they form the fastest growing segment of the population. At jump-off time the size of the population in ages 95+ was 3,700. We see from Figure 14 that the size of the population is expected to triple in the next 40 years. The summary statistics of the predictive distribution of the population in ages 95+, in 2030 and 2050, are (in thousands):

year	d_1	Q_1	Md	Q_3	d_9
2030	4.3	7.2	11.6	17.3	24.1
2050	4.8	12.1	27.2	51.3	83.6

Both the figure and the summary statistics indicate a very high level of uncertainty in forecasting, however. The uncertainty is partly due to migration, and partly due to the high level of uncertainty in the mortality of the oldest-old.

As women live longer than men, their share among the elderly is quite high. At the jump-off time it was 61% among those in ages 65+. As we have seen in Figure 2, the gap between female and male life expectancies is expected to narrow. This results in a declining share of females among the elderly, see Figure 15. Indeed, the summary statistics of the predictive distribution of the fraction of women in ages 65+, in 2030 and 2050, are:

year	d_1	Q_1	Md	Q_3	d_9
2030	0.53	0.54	0.55	0.57	0.58
2050	0.51	0.52	0.54	0.57	0.59

I.e., women's share does go down, although women will still remain the majority in this age bracket.

4.6. Beyond 2050

As our error estimates become increasingly untrustworthy when the forecast period is extended, we present here some summary data on the characteristics we have discussed above, for 2065:

characteristic	d_1	Q_1	Md	Q_3	d_9
life exp. female	84.8	86.3	88.0	89.8	91.4
life exp. male	78.7	81.0	83.6	86.0	88.1
total pop. (1000's)	3,256	3,824	4,640	5,625	6,529
age-dep. ratio	0.83	0.89	0.96	1.05	1.14
old age-dep. ratio	0.41	0.48	0.57	0.68	0.81
births (1000's)	17	27	42	67	100
pop. in age 0-6 (1000s)	126	197	301	468	681
pop. in age 7-18 (1000s)	259	381	552	800	1,119
pop. in age 19-64 (1000s)	1,589	1,959	2,385	2,870	3,367
pop. in age 65+ (1000s)	1,095	1,204	1,345	1,498	1,630
pop. in age 95+ (1000s)	4.8	13.2	29.9	58.9	97.2
fraction women in 95+	0.49	0.51	0.53	0.56	0.58

(Note that one cannot add population sizes in the columns and obtain the same fractile for the aggregate, because the sizes of the populations are less than perfectly correlated! For the same reason, one cannot get, e.g., the 9th decile for the old age-dependency ratio by dividing the 9th decile for the population in age 65+ by the 9th decile for the population in ages 19-64.) We conclude that the trends already visible become stronger as the forecast period is extended.

5. DISCUSSION

The forecast we have presented differs technically from conventional population forecasts in that it has been systematically phrased in probabilistic language. The goal has been to give a realistic indication of the level of uncertainty one can expect. In all cases we have attempted to provide an empirical basis for the assessment of uncertainty. Implicit in this is the assumption that future uncertainty is similar to that experienced in the past. This is the way the predictive distributions of this forecast should be interpreted.

Subjective elements have entered the specification of both the medians of the predictive distributions needed in cohort component forecasting, and in the specification of the spread around the medians. The effect of judgment has generally been to lower the level of uncertainty. For example, in the analysis of fertility and mortality the dramatic shocks of the 19th century were eliminated by the estimation method used.

The uncertainty of future migration is more difficult to estimate as reliably as that of fertility or mortality. Judgment necessarily has a greater role. However, even here the assessments have been grounded on empirically observed data.

References

Alho J.M. (1990b) Stochastic methods in population forecasting. *Int. J. Forecasting* 6, 521-530.

Alho J.M. (1998) A stochastic forecast of the population of Finland. *Reviews 1998/4.* Statistics Finland.

Alho J.M. (2000) A statistical look at Modeen's forecast of the population of Finland in 1934. *Yearbook of Population Research in Finland* XXXVI, 107-120.

Alho J.M. (2002) *Forecasting life expectancy: model choice and auxiliary data*. Paper prepared for the Committee on Mortality Forecasting, Swedish Social Insurance Board.

Alho J.M. and Spencer B.D. (1985) Uncertain population forecasting. *Journal of the American Statistical Association* 80, 306-314.

Alho J.M. and Spencer B.D. (1997) The practical specification of the expected error of population forecasts. *Journal of Official Statistics* 13, 203-225.

Box G.E.P. and Jenkins G.M. (1976) *Time Series Analysis, Revised Ed*. San Francisco: Holden-Day.

Burgdörfer F. (1932) *Volk ohne Jugend*. Berlin: Kurt Vowinkel.

Cannan E. (1895) The probability of cessation of the population of England and Wales during the next century. *The Economic Journal* 5, 505-515.

Population Projections 1973-2000. *Statistical Surveys N:o 52*, Central Statistical Office of Finland, Helsinki.

Cheeseman-Day J. (1993) Population projections of the United States by age, sex, race, and Hispanic origin: 1993 to 2050. *Current Population Reports, Series P-25, No. 1018.* U.S. Bureau of the Census.

DeGans H.A. (1999) *Population forecasting 1895-1945*. Kluwer, Dordrecht.

Draper D. (1995) Assessment and propagation of model uncertainty (with discussion). *Journal of the Royal Statistical Society* B, 45-97.

Fougstedt G. (1977) Trends and factors of fertility in Finland. *Commentationes Scientiarum Socialium 7/1977*. Societas Scientiarum Fennica, Helsinki.

Hjerppe, R. (1989) The Finnish economy 1860 - 1985, growth and structural change. Bank of Finland Publications, Studies on Finland's Economic Growth no. 13. Helsinki: Government Printing Centre.

Hoem J. (1973) Levels of error in population forecasts. *Artikler 61.* Oslo: Central Bureau of Statistics.

Hämäläinen H. (1987) *The accuracy of the regional population projections in Finland in the 1980's*. Paper presented in the Finnish-Soviet Seminar of Demography in Jyväskylä, June 1987.

I.N.E.D. (1976) *Natalité et politique démograhique*. Paris : Presses Universitaires de France.

Keilman N. (1990) *Uncertainty in national population forecasting: issues, backgrounds, analyses, recommendations*. Amsterdam: Swets and Zeitlinger.

Keilman N. (1997) Ex-post errors in official population forecasts in industrialized countries. *J. Official Statist*. 13, 245-277.

Keilman N., Pham D.Q. and Hetland A. (2002) *Norway's uncertain demographic future*. Social and Economic Studies 105. Oslo: Statistics Norway.

Lee R.D. (1974) Forecasting births in post-transition populations: stochastic renewal with serially correlated fertility. *Journal of the American Statistical Association* 69, 607-617.

Lee R.D. and Miller T. (2001) Estimating the performance of the Lee-Carter method for forecasting mortality. *Demography* 38, 537-549.

Lee R.D. and Tuljapurkar S. (1994) Stochastic population forecasts for the United States: beyond high, medium, and low. *Journal of the American Statistical Association* 89, 1175-1189.

Luther G. (1993) *Suomen tilastotoimen historia vuoteen 1970*. Helsinki: Tilastokeskus.

Modeen G. (1934a) *Suomen väkiluvun vastainen kehitys*. Tilastokatsauksia n:o 10. Helsinki: Tilastokeskus.

Modeen G. (1934b) Suomen väkiluvun tuleva kehitys ja sen taloudelliset seuraukset. *Kansantaloudellinen aikakauskirja*, VI, 351-378.

Modeen G. and Fougstedt G. (1938) *Laskelmia Suomen vastaisesta väestönkehityksestä*. Sarja VI, Väestötilastoa 89. Helsinki: Tilastokeskus.

Myrdal A. and Myrdal G. (1934) *Kris i befolkningsfrågan*. Stockholm: Bonniers.

National Research Council (2000) Beyond six billion. Washington D.C.: National Academy Press.

Statistische Reichsamt (1930) Volkszählung. *Statistik des Deutschen Reichs, Band 401,II.*, Berlin.

Talousneuvoston aluejaosto (1972) *Kokonaistaloudellinen alueellinen tarkastelukehikko, liiteosa 1-2*. Helsinki: Valtion painatuskeskus.

Törnqvist L. (1974) Näkökohdat, jotka ovat määränneet primääristen prognoosiolettamusten valinnan, in Hyppölä J., Tunkelo A. and Törnqvist L. (1947) *Suomen väestöä, sen uusiutumista ja tulevaa kehitystä koskevia laskelmia*. Tilastollisia tiedontantoja 38. Helsinki: Tilastokeskus.

Vaupel J.W. and Oeppen J. (2002) Broken limits to life expectancy. *Science*. To appear.

United Nations (2001) *World population prospects. The 2000 revision. Volume I: comprehensive tables*. New York: United Nations.

Whelpton P.K., Eldridge H.T. and Siegel J.S. (1947) *Forecasts of the population of the United states 1945-1975*. Washington D.C.: U.S. Bureau of the Census.

Väestöennusteryhmä (1973) *Väestöennusteiden laadinnan järjestäminen*. Helsinki: Valtioneuvoston kanslian julkaisu.

Wicksell S.D. (1934) Bidrag till den formella befolkningsteorien. Med tillämpningar på Norges befolkning. *Statsøkonomisk Tidsskrift* 1934, Hefte 1-2, 1-94.

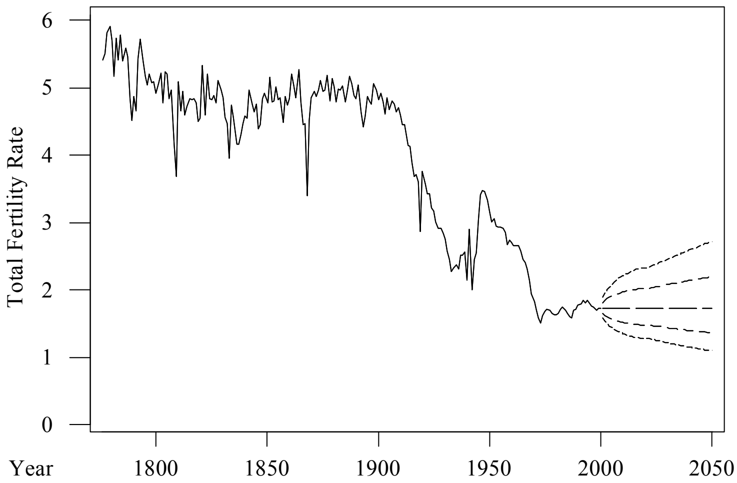

Figure 1. **The Total Fertility Rate of Finland in 1776-2000, and its Forecast in 2001-2050, with 50% and 80% Prediction Intervals.**

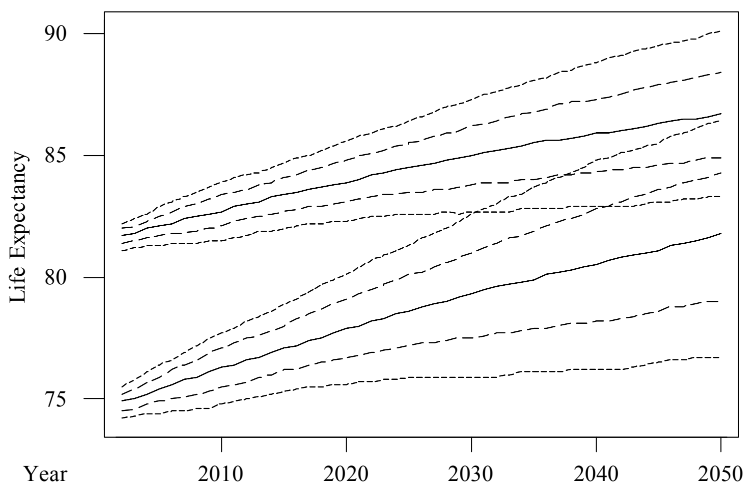

Figure 2. **Forecast of Life Expectancy, and its 50% and 80% Prediction Intervals, for Females (Upper Curves) and Males (Lower Curves), in 2001-2050.**

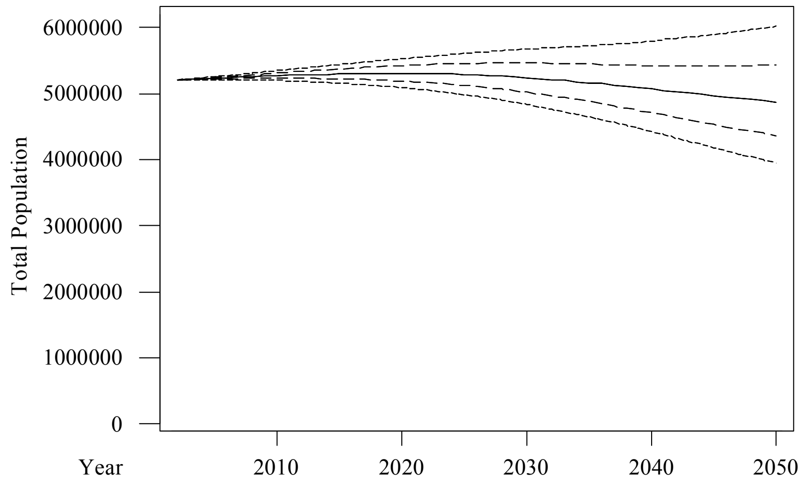

Figure 3. **Forecast of the Population of Finland in 2002-2050 with 50% and 80% Prediction Intervals.**

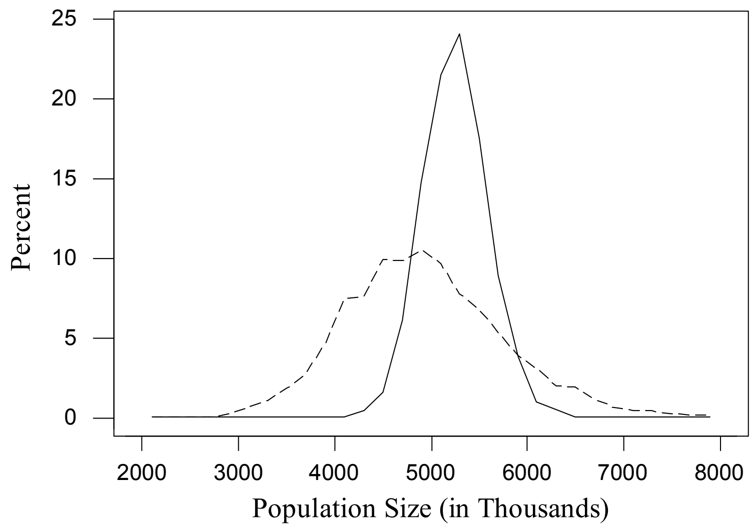

Figure 4. **Predictive Distribution of the Population in 2030 (Solid) and 2050 (Dashed).**

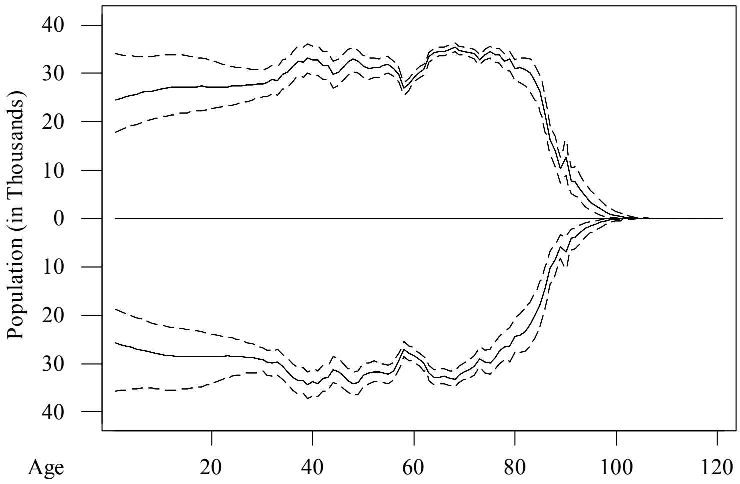

Figure 5. **Forecast of Population in 2030 by Age for Females (Top) and Males (Bottom): Median and 80% Prediction Intervals.**

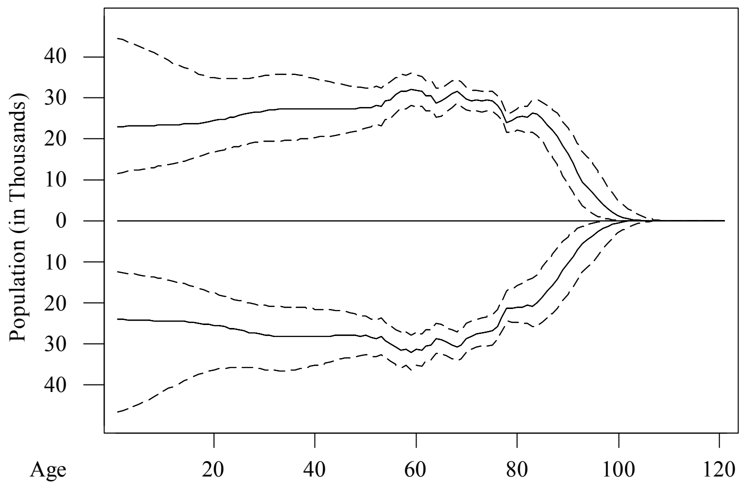

Figure 6. **Forecast of Population in 2050 by Age for Females (Top) and Males (Bottom): Median and 80% Prediction Intervals.**

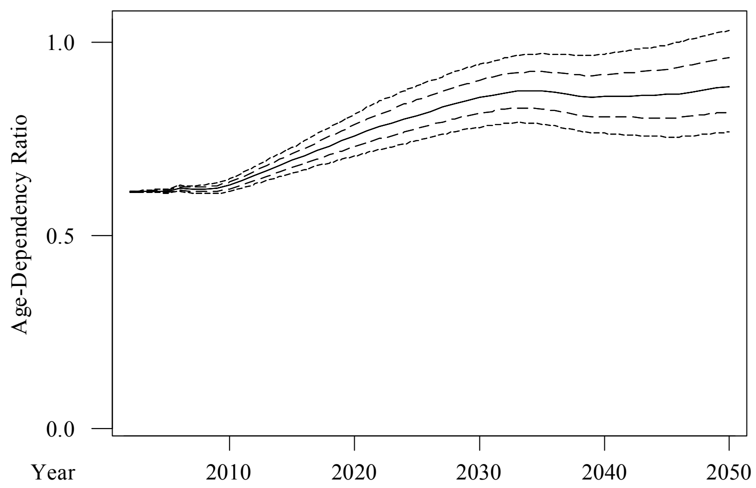

Figure 7. **Forecast of the Age-dependency Ratio in 2002-2050 with 50% and 80% Prediction Intervals.**

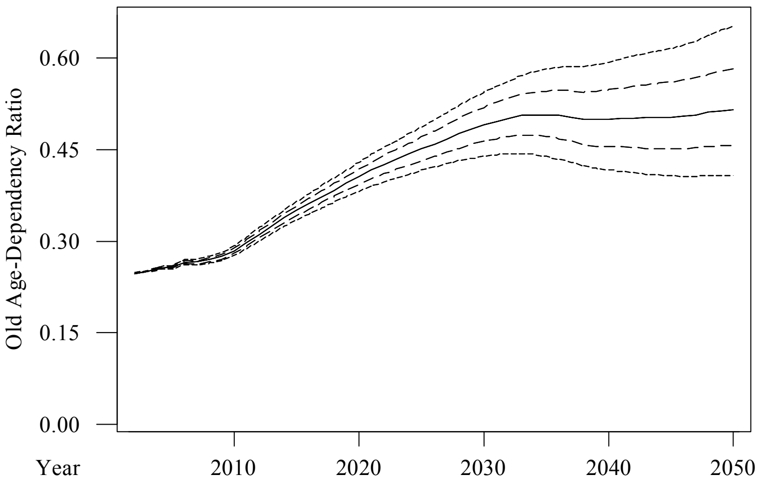

Figure 8. **Forecast of the Old Age-dependency Ratio in 2002-2050 with 50% and 80% Prediction Intervals.**

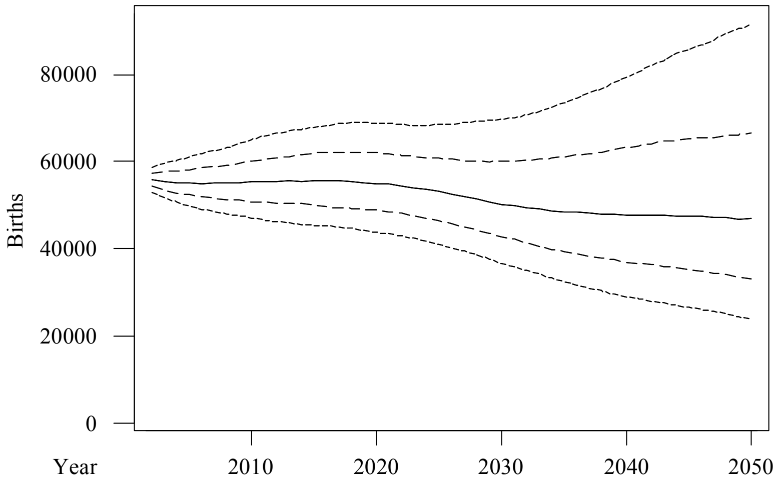

Figure 9. **Forecast of Births in 2002-2050 with 50% and 80% Prediction Intervals.**

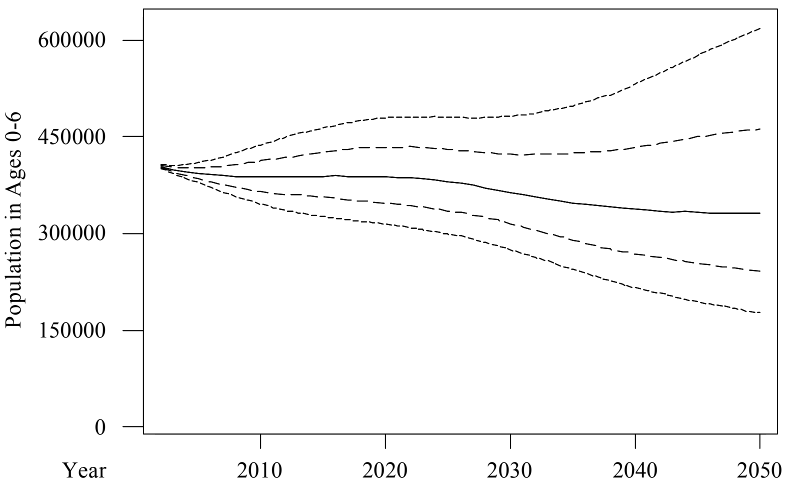

Figure 10. **Forecast of Population in Ages 0-6 in 2002-2050 with 50% and 80% Prediction Intervals.**

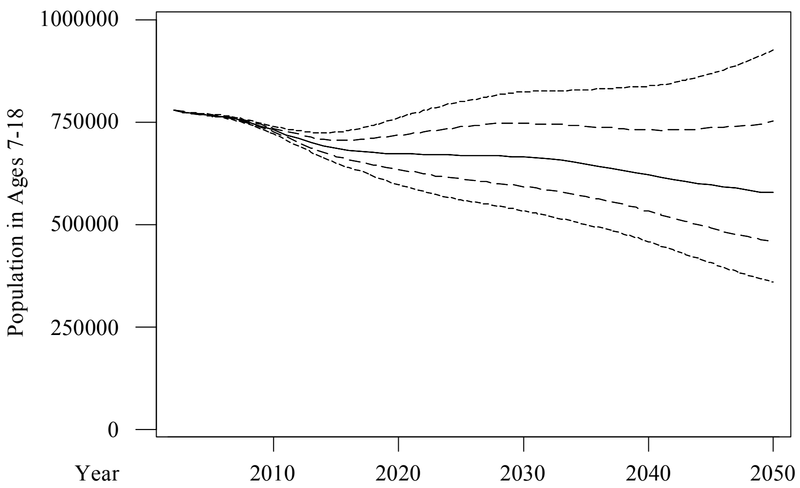

Figure 11. **Forecast of Population in Ages 7-18 in 2002-2050 with 50% and 80% Prediction Intervals.**

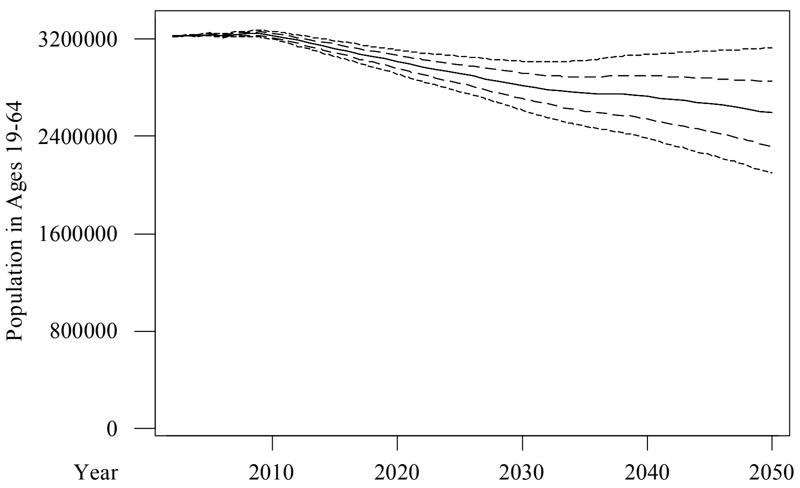

Figure 12. **Forecast of Population in Ages 19-64 in 2002-2050 with 50% and 80% Prediction Intervals.**

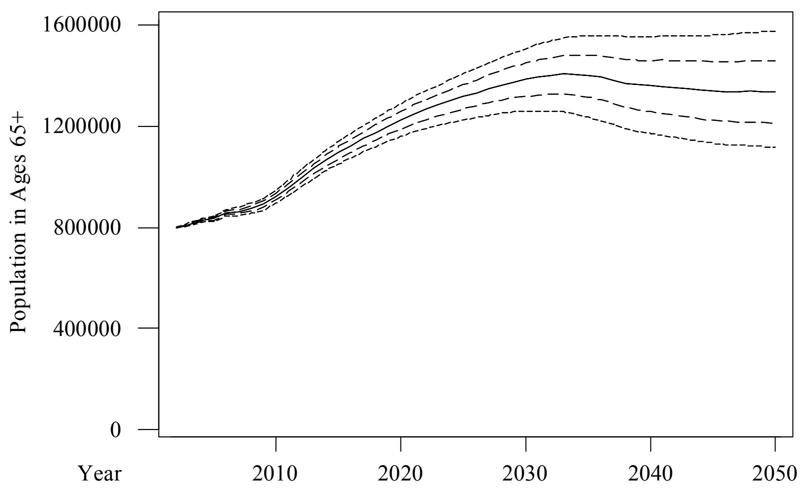

Figure 13. **Forecast of Population in Ages 65+ in 2002-2050 with 50% and 80% Prediction Intervals.**

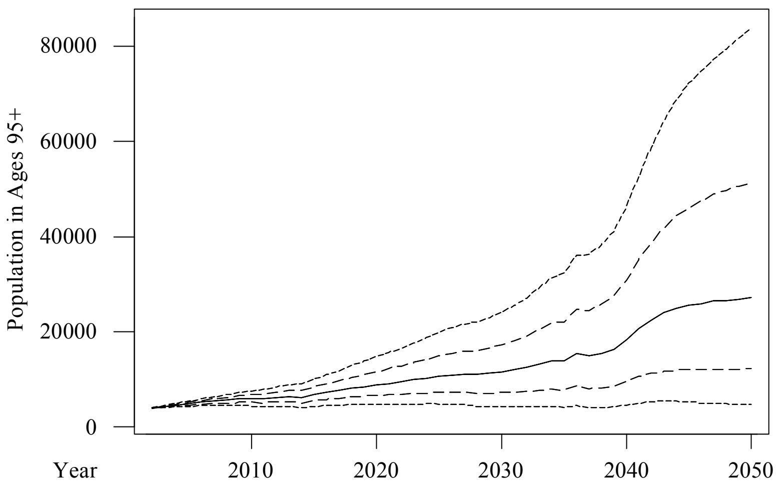

Figure 14. **Forecast of Population in Ages 95+ in 2002-2050 with 50% and 80% Prediction Intervals.**

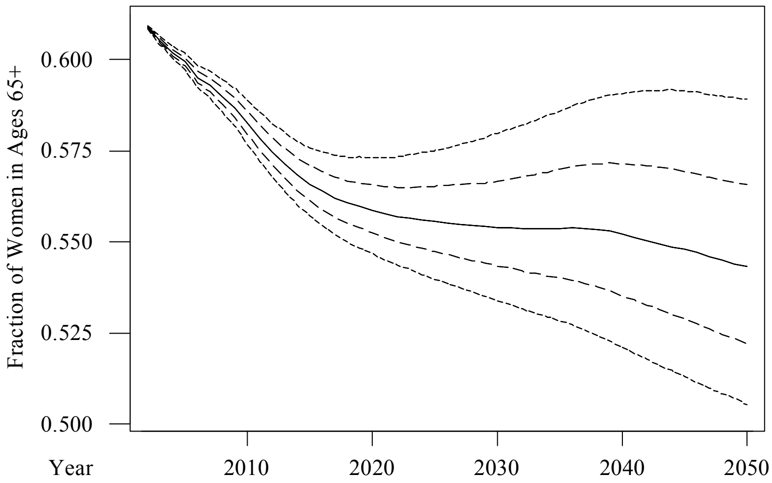

Figure 15. **Forecast of the Fraction of Women in Ages 65+ in 2002-2050 with 50% and 80% Prediction Intervals.**

Appendix. The Birth of the Baby-boom Generations in Finland.

The total numbers of (live) births in consecutive five-year periods, during 1925-1954, were as follows:

years	births
1925-1929	384,300
1930-1934	349,200
1935-1939	366,000
1940-1944	372,600
1945-1949	521,300
1950-1954	466,200

We see that the number of births reached a low during the years following the economic depression of the 1930's. After that there was a recovery, and during the five-year period that was most influenced by the war, the recovery continued: the total number of births was *higher* during the five years of war than during the previous five-year period of peace. It follows that the high number of births during 1945-1949 (with peak in 1947) cannot be attributed solely to a recovery of births that had earlier been prevented by the war.

Note that the above analysis does not mean that the war would not have had important timing effects. We see from Figure 1 that there was a zig-zag pattern caused by the winter war, the following short peace, and the subsequent continuation of war. In monthly data even finer effects of timing can be distinguished.

A more plausible explanation can potentially be given in terms of a longer term postponement caused by both the depression and the war. The validity of this can be investigated by studying completed cohort fertility. Figure A presents the sum of age-specific fertility rates in ages 15-40 for the birth cohorts born in 1905-1965.[3] Before analyzing the data, two technical remarks are in order.

First, due to the approximate method of estimation, the statistics of the cohort of 1905 are more accurately attributed to those born at the end of 1905, those of the cohort 1906 are more accurately attributed to those born at the end of 1906 etc. As discussed by Fougstedt (1977, p. 19), the approximations have a notable numerical effect for some birth cohorts hat were born at a time, when fertility was rapidly changing from month to month. The years 1918-1919, 1939-1940, and 1944-1946 are examples of this.

Second, for the last five cohorts the values have been forecasted by adding 0.16 to the cumulative fertility of ages 15-35. This is the difference observed for the last available cohort born in 1960. Given that the fertility in ages 40-49 has been approximately 0.05 during the 1940's and 0.01 recently, the cumulative sum for the ages 15-40 approximates cohort total fertility rate well.

Turning to Figure A, completed fertility presents a much smoother picture of the evolution of fertility than period fertility. This is to be expected, since fertility is heavily influenced by period factors that tend to compensate for each other for actual cohorts, over time. Nevertheless, completed fertility has changed during the period we are investigating. It started at level 2.3 for the cohort of 1905, and rose to a high of 2.7 for the cohort of 1919. As argued by Fougstedt (1977, p.18), the method of estimation has slightly exaggerated this value and decreased the low value of the previous year, so 2.6

[3] The author is grateful to Timo Nikander of Statistics Finland for providing these statistics.

may be closer to the actual maximum. From there, a decline to about 1.8 takes place. In other words, the increase during the early part of the period is about 0.3 children, and the subsequent decline is about 0.8 children, or 31%. From an even longer-term perspective the baby-boom still appears as a reversal of a declining trend that started in the late 1800's and continued after the 1950's.

In thinking about the possible reasons for a reversal of a long-term decline, it seems useful to look at other countries, as well. Sweden did not participate in war, but had a baby-boom that peaked in 1945, and a smaller peak in 1964. Great Britain and Belgium had lesser peaks in 1947-1948 and a bigger one in 1964. France and the Netherlands had higher peaks in 1946-1947 and a lesser one in 1964. The United States and Canada had major peaks in 1957 and 1960, respectively. (I.N.E.D. 1976, pp. 46-54) Clearly, Finland's experience does not exactly match that of any of the other countries mentioned. However, in terms of completed cohort fertility all the countries share the same feature: a temporary reversal of a long time declining trend took place.

The conclusions one can draw from the above are dissatisfying. We have ample evidence that on a cohort level a temporary reversal took place. This corresponds to *underdamped* systems in control theory (Box and Jenkins 1976, p. 344), but merely finding an analogue for the phenomenon hardly explains why it happened. Moreover, from the perspective of forecasting births it is the period fertility that counts, because it determines the annual births. This, of course is at the root of the problems we are experiencing now, when the baby-boom generations are about to retire.

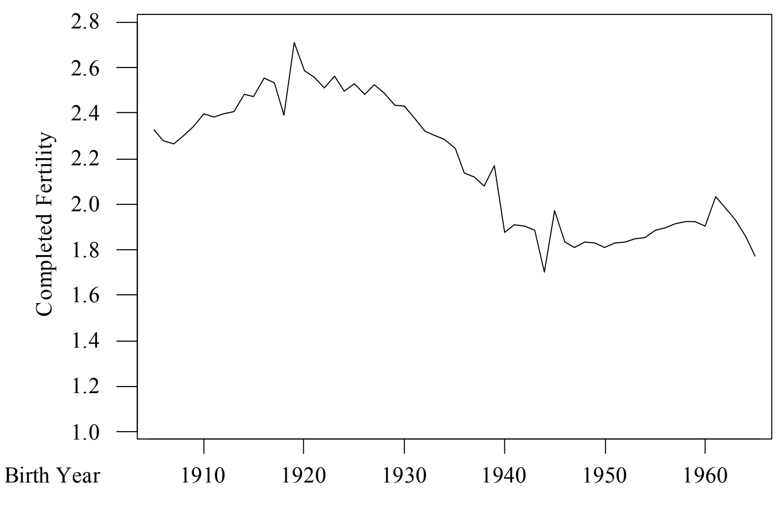

***Figure A*. Approximate Completed Fertility for Birth Cohorts Born in 1905-1965.**

E L I N K E I N O E L Ä M Ä N   T U T K I M U S L A I T O S   (ETLA)
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B, FIN-00120 HELSINKI

Puh./Tel. (09) 609 900
Int. 358-9-609 900
http://www.etla.fi

Telefax (09) 601753
Int. 358-9-601 753

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 798 PETRI BÖCKERMAN, Understanding Regional Productivity in a Nordic Welfare State: Does ICT Matter? 22.03.2002. 20 p.

No 799 JYRKI ALI-YRKKÖ – RAINE HERMANS, Nokia Suomen innovaatiojärjestelmässä. 11.04.2002. 35 S.

No 800 JUHA FORSSTRÖM – JUHA HONKATUKIA, Energiaverotuksen kehittämistarpeet Kioton pöytäkirjan toteuttamiseksi. 26.03.2002. 31 s.

No 801 JUHA FORSSTRÖM – JUHA HONKATUKIA – ANTTI LEHTILÄ, Maakaasun hinnan vaikutukset ilmastopolitiikan kustannuksiin. 27.03.2002. 26 s.

No 802 ELINA RISSANEN, Reflections of Globalisation – The Influence of Economic Factors on the International Development Assistance Allocations. 09.04.2002. 19 p.

No 803 ELINA RISSANEN, Foreign Aid Policies – Comparing Finland and Ireland. 09.04.2002. 22 p.

No 804 RAIMO LOVIO, Suomalaisten monikansallisten yritysten T&K-toiminnan kansainvälistyminen – perustietoja ja kysymyksenasetteluja. 11.04.2002. 22 s.

No 805 ARI HYYTINEN – MIKA PAJARINEN, External Finance, Firm Growth and The Benefits of Information Disclosure: Evidence from Finland. 24.04.2002. 40 p.

No 806 MATTI LAMPINEN, Uusi talous – tapaustutkimuksia suomalaisyrityksissä. 25.04.2002. 13 s.

No 807 OLAVI RANTALA, Kansantalouden ja toimialojen pitkän ajan kasvu – Vertailua ilmastostrategian BAU-skenaarioon. 06.05.2002. 11 s.

No 808 JUSSI V. KOIVISTO – MATTI LAMPINEN, Comparative Analysis of Finno-Sinic Management: In-Depth Study of the Dimensions of Cross-Cultural Friction and Synergy in Management of Finnish Firms in Hong Kong. 22.05.2002. 18 p.

No 809 JUKKA LIIKANEN – PAUL STONEMAN – OTTO TOIVANEN, Intergenerational Effects in the Diffusion of New Technology: The Case of Mobile Phones. 22.05.2002. 16 p.

No 810 TIMO KAISANLAHTI, Minority Shareholders in the Finnish System of Corporate Governance. 10.06.2002. 71 p.

No 811 JYRKI ALI-YRKKÖ – RAINE HERMANS, Nokia in the Finnish Innovation System. 19.06.2002. 35 p.

No 812 ARI HYYTINEN – MIKA PAJARINEN, Small Business Finance in Finland. A Descriptive Study. 25.06.2002. 59 p.

No 813 ARI HYYTINEN – MIKA PAJARINEN, Financing of Technology-Intensive Small Businesses: Some Evidence on the Uniqueness of the ICT Industry. 29.07.2002. 22 p.

No 814 ARI HYYTINEN – OTTO TOIVANEN, Does Distance Matter in Deposit Supply? A Bank-level Study. 30.07.2002. 14 p.

No 815 ERKKI KOSKELA – MARKKU OLLIKAINEN, Optimal Forest Taxation under Private and Social Amenity Valuation. 13.08.2002. 23 p.

No 816 TARMO VALKONEN, Demographic Uncertainty and Taxes. 21.08.2002. 20 p.

No 817 ANTTI KAUHANEN – HANNU PIEKKOLA, Profit Sharing in Finland: Earnings and Productivity Effects. 03.09.2002. 18 p.

No 818 ISMO LINNOSMAA – RAINE HERMANS – TARU KARHUNEN, Price-Cost Margin in the Pharmaceutical Industry: Empirical Evidence from Finland. 03.09.2002. 20 p.

No 819 RAINE HERMANS – TERTTU LUUKKONEN, Findings of the ETLA Survey on Finnish Biotechnology Firms. 10.09.2002. 30 p.

No 820 ARI HYYTINEN – OTTO TOIVANEN, Do Financial Constraints Hold Back Innovation and Growth? Evidence on the Role of Public Policy. 17.09.2002. 31 p.

No 821 KARI E.O. ALHO, Stbilization Inside and Outside EMU. 27.09.2002. 20 p.

No 822 HANNU PIEKKOLA, From Creative Destruction to Human Capital Growth: Wage Dispersion Effects in Finland. 27.09.2002. 20 p.

No 823 ARI HYYTINEN – OTTO TOIVANEN, Misuse and Non-use of Information Acquisition Technologies in Banking. 11.10.2002. 14 p.

No 824 HELI KOSKI – TOBIAS KRETSCHMER, Entry, Standards and Competition: Firm Strategies and The Diffusion of Mobile Telephony. 14.10.2002. 36 p.

No 825 PEKKA SULAMAA – MIKA WIDGRÉN, EU-Enlargement and the Opening of Russia: Lessons from the GTAP Reference Model. 15.10.2002. 24 p.

No 826 JUHA M. ALHO, The Population of Finland in 2050 and Beyond. 11.11.2002. 28 p.

No 827 JUKKA JALAVA, The Production and Use of ICT in Finland, 1975-2001. 21.10.2002. 23 p.

No 828 ARI HYYTINEN – TUOMAS TAKALO, Enhancing Bank Transparency: A Re-assessment. 23.10.2002. 23 p.

No 829 REIJO MANKINEN – PETRI ROUVINEN – PEKKA YLÄ-ANTTILA, Palveluiden tuottavuus – kilpailu ja teknologia muuttavat rakenteita. 31.10.2002. 49 s.

No 830 PEKKA MANNONEN, The Strategic Response of Banks to an Exogenous Positive Information Shock in the Credit Markets. 31.10.2002. 16 p.

No 831 JYRKI ALI-YRKKÖ – PEKKA YLÄ-ANTTILA, Pääkonttorien sijainti, kansainvälistyminen ja verotus. 05.11.2002. 38 s.

No 832 ARI HYYTINEN – LOTTA VÄÄNÄNEN, Government Funding of Small and Medium-sized Enterprises in Finland. 06.11.2002. 50 p.