A reconstruction of the population of Scania 1650-1760

Bengtsson, Tommy; Oeppen, Jim

1993

Document Version:
Other version

Link to publication

Citation for published version (APA):
Bengtsson, T., & Oeppen, J. (1993). *A reconstruction of the population of Scania 1650-1760*. (Lund Papers in Economic History. Population Economics; No. 1993:32).

Total number of authors:
2

Lund Papers in Economic History are published by the Department of Economic History, Lund University, Sweden. This series replaces the former series under the title *Meddelande från Ekonomisk-historiska institutionen, Lunds universitet*. The change of name reflects the orientation of the series towards an international readership. The series is multilingual, but the majority of the working papers appear in English.

Lund Papers in Economic History include papers in the following topic areas:

- *General Issues*
- *Development Economics*
- *Education and the Labour Market*
- *Population Economics*

The division into topic areas is due to the wide range covered by the Department's researchers, and it provides the basis for topic-orientated distribution.

Lund Papers in Economic History are published as occassion arises, not at fixed intervals. Single back numbers cost SEK 35 each, and a subscription for 5 issues costs SEK 125. Subscriptions are accepted for the series as a whole or for working papers in individual topic areas.

Lund Papers in Economic History

No. 32, 1993

Population Economics

ISRN LUSADG-SAEH-P--93/32--SE+23P

A Reconstruction of the Population of Scania 1650-1760

Tommy Bengtsson & Jim Oeppen

DEPARTMENT OF ECONOMIC HISTORY, LUND UNIVERSITY

Q(5)

Lund Papers in Economic History
ISSN 1101-346X

Printed in Lund, Sweden, 1993
Reprocentralen
Lund University

Orders and Subscriptions
Department of Economic History, Lund University
Mail address: P.O. Box 7083, S-220 07 Lund, Sweden
Telephone: +46 46 107475
Telefax: +46 46 131585

ISRN LUSADG-SAEH-P--93/32--SE+23P

A Reconstruction of the Population of Scania 1650-1760

*Tommy Bengtsson & Jim Oeppen**

The interest in the demographic conditions prevailing during the industrial and agricultural revolutions, in combination with the lack of population records for most countries for that period, has led to a focus on Sweden which has regular censuses from 1749 onwards. Sweden has therefore been used as a model for the theory of the demographic transition.[1] This theory, which was first formulated in the 1940s has been heavily criticised in the past decades. One such critique is that reconstructions of the population in England and France have shown that the demographic transition could follow very different routes.[2] The basic assumption in the theory of the demographic transition, that mortality and fertility were at a high level prior to the transition, although fluctuating violently, has proved to be wrong for England and France. For both countries there exist secular changes in mortality and fertility prior to the demographic transition. For England we also know that the mortality level varied indepen-

* Tommy Bengtsson is Associate Professor in Economic History at Lund University and Jim Oeppen is Senior Research Associate of the Cambridge Group for the History of Population and Social Structure. This paper has been written within the research project "Life Events in a peasant society in transition" financed by the Swedish Council for Research in the Humanities and Social Sciences.

1 Davis, K (1945), "The World Demographic Transition", *Annals of the American Academy of Political and Social Science* 237:1-11

2 Wrigley, E A and Schofield, R S (1981), *The Population History of England 1541-1871 - A Reconstruction*, London, and Henry, L and Blayo, Y (1975), "La Population de la France de 1740 à 1860", *Population*, xxx, pp 403-42.

dently of economic conditions while the levels of nuptiality and fertility followed the economy.

Thus we know that England and Sweden took different routes during the demographic transition. The question arises whether their ways were different before as well? Did the levels of mortality and the fertility in Sweden, as in England, vary before the transition started or were they stable as the theory of demographic transition assumed? To answer this question we need information further back in time than the first census in 1749. We need an alternative source of information and an appropriate method for non-census data.

One way will be to undertake family reconstitutions. This method is, however, very labour intensive. It is therefore difficult to cover more than fractions of a nation's population. Furthermore, for most countries one is able to reconstitute only a small proportion of the families. We will therefore argue that the best source is the aggregated parish records of births and deaths and that the most appropriate method is Generalised Inverse Projection.[3] At this stage we will not reconstruct the entire population of Sweden, but only the population of the county of Scania in the southernmost part of the country. For Scania, we have population records back to the middle of the 17th century. The reconstruction of the size of population in Scania not only serves the purpose of comparison with England and France but also can be compared with the results of family reconstitutions of a group of parishes in Scania.[4]

Before examining GIP and reconstructing Scania's population back to 1650, we will make a brief summary of two other back-projections of the population size in Sweden made by Eli Heckscher and Lars Widén respectively.

3 The method was developed by Jim Oeppen at the ESRC Cambridge Group for the History of Population and Social Structure, England. See Oeppen, J (1993), "Back-Projection and Inverse Projection: members of a wider class of constrained projection models", *Population Studies*, Vol 47, pp 245-267.

4 The project is described in Bengtsson, T and Lundh, C (1990), "Life Events in a Peasant Society in Transition. Causes of Population Change in Scania, 1660-1860". *Lund Paper in Economic History*, No 1, Lund.

Back Projections

The method Heckscher used when estimating the Swedish population back to 1721 was to subtract the natural population increase from the population size in 1748 or 1749 assuming a closed population, i.e. net migration was zero.[5] He used data on numbers of births and deaths from 1721 to 1748, in practice the number that were baptised and buried. The data were collected annually from 1736 and as a retrospective investigation for the period 1721 to 1735. The tables still available do not cover the entire country, but only 9 counties and Stockholm, comprising a little more than 40 per cent of the total population in 1750. For some counties the figures for 1721 to 1735 are not reported annually but only as a summary. For others, the numbers of deaths are reported by sex and age-group. Heckscher, however, only made use of the total numbers of deaths. On the basis of the numbers of births and deaths and estimated population size, the crude birth and crude death rates were calculated. The major results are that the mortality level was considerably higher in the period 1736 to 1748 than earlier, while fertility was more stable, indicating a substantial population increase between 1721 and 1735. Heckscher found support for these findings in his calculations of population increase from taxation records.[6] He argued that the low mortality rate during the 1720s and early 1730s was a result of excess mortality during the great Nordic wars that ended in 1721. This explanation has met criticism.[7]

Heckscher's back projection of the Swedish population size is an example of the use of existing information on births and deaths, but he did not make the fullest use of the information about the population in 1748/49. He only used the size of the total population in that year and not its age-structure. In a retrojection back to 1700, Widén made use of information about the age structure during the period 1750-1800 and age-specific fertility in 1751-1774, as well as the same annual data

5 Heckscher, E (1936), "Sveriges befolkning från det stora nordiska krigets slut tilltabellverkets början (1729-1750)", p 261-2, in Heckscher, E (1936), *Ekonomisk-historiska studier*, Stockholm. The assumption of zero net migration is justified by the fact that the net out-migration during the second half of the 18th century was estimated to be only 1500 per year. Furthermore, for the period 1721 to 1748, migration was not discussed in any sources known to Heckscher.

6 Heckscher, E (1933), "En mätare på svenska folkets välståndsutveckling. Den mantalsskrivnabefolkningen 1634-1820". *Svensk Historisk Tidskrift* 1933, pp 365.

7 See for example Gille, H (1949), "The Demgraphic History of Northern Europe in the 18th Century", *Population Studies*, vol III, pp 3-65.

on births and deaths for the period 1736 to 1748 that Heckscher used.[8] Using stable population theory he projected the population backwards to 1700, correcting for losses during the battle in Narva in 1700 and the battle in Poltava in 1709. The reason Widén did not make use of the figures on births and deaths for the period 1721-1735 was that he thought that they were not complete. Not surprisingly, Widén estimated approximately stable birth and death rates for the entire period. Thus, the method used by Widén makes use of the inherent information in the age-structure as well as information about numbers of births and deaths for part of the period. The problem remains that he was forced to make strong assumptions about survival and fertility for the period 1700 to 1735, when he had no information on births and deaths.

The general situation is that we have no aggregate information on numbers of births and deaths before the first census or, as for Sweden as a whole, only for a short period before the first census. Thus, we must make use of records at low aggregate levels, generally at parish level and primarily on baptisms, burials and marriages. These records can either be used directly or as a source for family reconstitution. From family reconstitution, which is very time consuming, we will only cover a small sample of the population and for this sample we are unable to get information on the entire population, since one is unable to reconstruct all families. The possibilities seem, however, to be better for some countries, like Canada and Sweden, but the method is very labour intensive even using computerised techniques.

Estimating population size, or change, from other sources, such as taxation records which Heckscher tried, is less time consuming, but even more difficult for two reasons. The information about family members and servants may be incomplete and some households that are too poor to pay any tax are not always reported in the taxation records. Similar but better sources than the taxation records are the catechism examination records. Every year the clergyman had to examine all the members of his parish of age 10 years and above on their ability to read and understand the catechism. In the records that were kept, information about the entire family and servants is given. However, as with taxation records, some persons might be missed. Children who die shortly after birth are seldom registered, so these

8 Widén, L (1976), "A retrojection back to 1700." Appendix II in Hofsten, E and Lundström, H (1976), *Swedish Population History. Main trends from 1750 to 1970.* Stockholm.

records have to be supplemented with information on births and deaths from the parish registers. Another disadvantage is that the catechism records are less common than the birth and death books for the 17th and 18th centuries.[9] Comparisons between estimations based on the catechism records and the taxation records were made by Friberg, who also discussed earlier attempts to estimate population size.[10] His results underline the problems in estimating population size from taxation and catechism records.

Collecting information on totals numbers of baptisms and burials takes only a fraction of the time it takes to undertake reconstitutions. A major disadvantage is that we cannot calculate birth and death rates as we do not know the population size. For analyses of short term changes this is not a major problem, since population size changes slowly compared to the changes in births and deaths. However, when it comes to long term changes in fertility and mortality we are unable to proceed without knowing the size of the population.

In 1974, Lee published a description of 'inverse projection', which was an attempt to satisfy the need for estimates of fertility and mortality when only the flows of births and deaths are known.[11] If it is assumed that a population is closed to migration, a series of population totals can always be calculated if the intervening totals of births and deaths are known. Lee's innovation was to show that, given a starting census and information on age-structure and the typical patterns of age-specific fertility and mortality rates, there is a unique series of fertility and mortality levels which exactly reproduces the observed series of population, birth and death totals and that successive age structures can be calculated from this information. In a later work, he showed that the zero-migration assumption could be replaced by specifying epochs with given
migration rates.[12]

9 For an overview of the availability of the catechism records, see Lundh, C and Bengtsson, T (1989), "Familjerekonstruktion på svenskt kyrkoboksmaterial. Problem och möjligheter", *Meddelande frånEhonomisk-historiska institutionen*, Lunds universitet, Nr 59, Lund.

10 Friberg, N (1953) *Dalarnas befolkning på 1600-talet*, Stockholm.

11 Lee, R D (1974), "Estimating Series of Vital Rates and Structures from Baptisms and Burials: A New Technique, with Application to Pre-Industrial England". *Population Studies*, 28(3), p 311-342.

12 Lee, R D (1985), "Inverse-Projection and Back-Projection: A Critical Appraisal, and Comparative Results for England, 1539 to 1871". *Population Studies*, vol 39, p 233-238.

'Back-projection', the major method used in 1981 in Wrigley and Schofield's reconstruction of the English Population back to 1541, was an attempt to free inverse projection from the necessity of specifying a fixed migration rate and the size and age-structure of the starting population. The method has since been developed by Oeppen.[13] Rather than comparing back- and inverse-projection methods in detail, Oeppen defined a wider class of Generalised Inverse Projection models and related both of them to that structure. For the estimation of the Scanian population from 1650 onwards we have used the GIP model.

The following inputs are needed for the GIP model in its minimal form.[14] A terminal census, which gives the population size and age structure at the end of the period, the annual numbers of births and deaths and age-specific fertility, mortality and net-migration rates representative of the period. Annual numbers of marriages, if available, are used for calculation of nuptiality rates. The age-group interval, which should equal the time unit for numbers of births and deaths, is usually 5 years. The sexes may be combined or treated separately.

The GIP model assumes that the age-specific pattern of deaths and net-migration in any period can be derived from the 'typical' age-specific schedules by raising or lowering their intensity according to the value of a single parameter for each process. Thus, if there are T five-year time periods in the data, GIP needs to estimate 2T parameters to control migration and mortality. In this way, the levels of mortality and net migration are allowed to change between each time unit, but their shapes are derived from the input schedules.

GIP estimates a 'census' for each period by projecting the given births, whilst making the additions or decrements for migration and mortality that are specified by the estimated 2T parameters.

The optimisation algorithm adjusts the parameters until it finds a population system that best fits the known data according to a specified criterion, rather than exactly matching the T observed totals of deaths and the K age-groups in the terminal census. We have to recognise that, no matter how reasonable our assumptions, no population will behave exactly as specified by the model schedules, due to the errors and random fluctuations that may occur in the counts of annual

13 See note 3.
14 The method is explained in detail in Oeppen, J see note 3.

events, the age-structure at the end of the period or in the age-specific rates.

A problem remains because we want to estimate censuses from the beginning of the data, but the censuses for the first 100 years contain a diminishing number of age-groups whose birth-cohort sizes are unknown. GIP assumes, if no other information can be provided, that the population system was evolving in a stable fashion before the period of the given data. With this assumption, there is then sufficient information to project backwards from the complete cohorts to those birth-totals are unknown. In long runs of data, the results can be quite sensitive to this assumption, but the data series for Scania is short and therefore severely constrained by the terminal census.[15]

Even if the model assumptions are met, the uniqueness of the reconstructed population depends on the relationship between the number of parameters 2T and the minimum number of 'targets' T+K. In most applications, 2T will be greater than T+K, which means that there is no unique solution to the estimation problem and that any reconstruction of the population is simply selecting one possibility from an infinity of solutions. The best policy is to increase the number of data targets but, if this is not possible, one way to proceed is to apply constraints to the parameter values, which effectively reduces the number to be estimated. In this paper we have decided to estimate populations which match the observed data, but also keep the time-series of net-migration parameters as smooth as possible. Smoothing is one of the fundamental elements of data-reconstruction techniques and is widely applied.

The fit to the data, the additional constraints on migration and the balance between these conflicting aims are brought together in an objective function, which is minimised using a standard numerical procedure. The objective is to

$$\underline{\min}\ F = \sum_{t=0}^{T} [(D_t - \hat{D}_t)/D_t]^2 + \sum_{a=0}^{K} [(N_{a,T} - \hat{N}_{a,T})/N_{a,T}]^2 + \lambda \sum_{t=0}^{T-1} (\hat{\beta}_{t+1} - \hat{\beta}_t)^2$$

$$\underline{\text{subject to}} \quad 0 < \alpha_t < \infty$$

$$\underline{\text{and}} \quad -\infty < \beta_t < 1$$

where alpha and beta are the mortality and migration parameters, D is

15 The dangers of this assumption for long data-series are explored in the chapters by Oeppen and Lee in Reher, D S and Schofield, R S (eds) (1993) *Old and New Methods in Historical Demography*. Oxford.

a five-year total of deaths, N is a census age-group and lambda balances fidelity to the data with fidelity to the assumption of smooth migration.

Additional data constraints or assumptions can be added to this objective function and the individual elements could be weighted. For example, in this study, we experiment with using the total number of persons aged 15 and over in 1699/1700 as an additional target.

The first two terms of the equation penalises any discrepancy between the data and the estimates, whereas the third term penalises any departure from smoothness in the net-migration parameters. The lambda parameter is an explicit reflection of the relative importance attached to the two prescriptions. Thus, depending on prior knowledge about the quality of the data and on net-migration, as well as on the length of the time period, we can choose whether the net-migration differentials should be regarded as an important factor or not when calculating the population for each unit of time. If we regard migration as of lesser importance, then lambda should be small in the above equation. On the other hand, if smooth migration is regarded as important as discrepancies in deaths and age-structure then lambda should be set to 1.

Selecting a value for lambda presents a problem. The user of GIP is forced to choose between fidelity to the data or, recognising that there may be problems with the data, the assumptions or both, that a smoothed reconstruction may be preferable to one which automatically transfers these problems to the results. As with the smoothing of time-series in general, there are few formal procedures available. Testing GIP, in cases where there are long runs of known populations for comparison, showed that optimum values for lambda lay between 0.2 and 1.0.[16] As we will demonstrate, it would appear that the short length of the Scania data series makes the population estimates insensitive to changes in lambda over this range.

The objective function represents a flexible method for combining data and assumptions quantitatively. It also allows the GIP user to assess the contribution to the fit of the model of each individual data point and assumption in the equation. As we shall see below, the errors or discrepancies that are summed in F are the major source of internal diagnostic information on the behaviour of the model.

16 See figure 5(a) in Oeppen, J (1993), "Back-Projection and Inverse Projection: members of a wider class of constrained projection models", *Population Studies*, Vol 47.

The GIP method has not only been used for reconstructing the English population back to 1541, but has also been tested on Norway, Finland and Sweden, as well as on regional and urban populations. These tests have been done for periods when data on population sizes exist, so that a comparison can be made. The results are very good, even in the case of Stockholm, a city with high migration. At the moment the method is being used in a number of studies of small populations in Europe, such as Amsterdam.[17]

The Data

The main objective is to estimate the size of the population in a sample of parishes in Scania, in southern Sweden, from 1650 to when the censuses start in 1749. A sample of about 80 parishes, covering approximately a quarter of the population in Scania, has been used in a number of previous studies - mainly to analyse the secular mortality decline and short-term population movements.[18] These studies have been based largely on the population tables from 1749 onwards containing aggregated information.

For 29 of the parishes in the sample we have parish records back to the end of the 17th century and for some of them back to the 1640s. In addition we have data for parishes for which family reconstitutions are being undertaken as well as for some other parishes.[19] Taken together, we have annual data on births and deaths, as well as marriages, for 43 parishes back to the 17th century, but only for 9 back to 1650. Mainly because of difficulties in reading the original documents, data are missing for some years in a number of the parishes. In these cases we have set the numbers of events equal to the average for the five year period. This has been done for only a few per cent of the total number of data points. Parishes with larger gaps were excluded from the sample. As the number of parishes in observation is increasing over time, the figures for earlier periods were inflated in an

17 Oeppen, J and van Leeuwen, M H D (1963), "Reconstructing the demographic regime of Amsterdam, 1681-1920", *Economic and Social History in the Netherlands*, Vol V, pp 61-102.

18 The sample has been used in a large number of studies. For an overview, see Fridlizius, G (1990), "The Growth of a Population Research Group in Economic History", Scandinavian Economic History Review, Vol XXXVIII, No:2.

19 These data are used in a the research project referred to in note 4.

attempt to estimate the level of events that would have been experienced by all 43 parishes. This problem was also faced in the reconstruction of the English population by Wrigley and Schofield.

As a terminal point for the reconstruction we have chosen 1760, although census data do exist for 1749 and onwards. The reason is that we believe that the first censuses may contain errors owing to their novelty. Thus the population in five year age-groups is taken from 1760. This is also the case for the age-specific fertility and mortality rates, which do not refer to the sample but to the whole of Scania (see Appendix). The reason we have chosen rates from the entire population rather than the sample is that it seemed preferable to use a larger population to minimise random variation. The age-specific migration rates refer to Sweden during the 1860s. They are the same as those used in the reconstruction of the English population.[20] It is generally believed that the age-profile of migration was quite stable until the end of the 19th century when the share of family migrants increased.[21] Whether this holds true for periods immediately after mortality shocks or not, we simply don't know. One would, however, expect a temporary change in the composition of migrants towards families after a mortality shock since relatively more adults are dying.

The information about births, deaths and marriages from 1650 to 1760 and the population in five-year age-groups in 1760 constitute our first, and major, data set. A second data set consisting of the 29 parishes from the original sample has been set up for the period 1750 to 1850. This was done in order to test the consistency of the GIP method and the data, in particular the fertility, mortality and migrations schedules used. We also wished to test if the sample size is large enough. The terminal age-structure refers to 1850. The migration rates are the same as for the earlier data set and the age-specific rates refer to the whole of Scania in 1850. Thus, the data set is very similar, particularly with respect to sample size, to the first one, but it covers a period for which we have full information about population size. To get the best possible estimation of the population prior to 1750 one should use the entire data set, i.e. data for 1650 to 1850. Since adding information on age-structure for all parishes for the period 1750 to 1850 is quite labour intensive and since we would like to use the method for even larger samples, eventually for the whole of Sweden, we have chosen to try to get good estimations using data from the middle of the 18th century and backwards.

20 Hofsten, E and Lundström, H (1976), *Swedish Population History. Main trends from 1750 to 1970.* Stockholm.

21 See for example Bengtsson, T (1987), "Migration och löner. Tillämpning av Todaros migrationsteori på 1800-talets svenska urbanisering", p 30-33, in *Ekonomisk-Historiska Vingslag*.Lund.

Concerning data quality, the major drawback is that the data are on baptism and burials and not births and deaths. However, the period between birth and baptism was very short until well into the 19th century.[22] Before then the baptism had to take place within eight days after the birth according to the church law from 1686. So loss of information about births due to the fact that a child dies before it was baptised is likely to be greater for our second data set than our first. The practice of emergency baptism by the midwife or other persons in the case when the child was likely to die also closes the gap between the number of baptisms and births. However, under-registration may still occur since both baptisms and burials were costly. This seems particularly the case for less well-off people.[23] However, we still do not know for which time periods and for which parishes the problem arises. Thus, we have to keep in mind that our figures for births and deaths are likely to be a little low.

Estimates of Population from 1750 to 1850

This is a period with a substantial population increase but also a period with mortality crises. Both in some years in the 1750s and in the 1770s the mortality is 4-5 times higher than normal, the highest in the 1770s. The periods contains several mortality crises but none due to plague - the last one in Scandinavia was in 1711. The population

22 Bengtsson, T och Lundh, C (1989), "Familjerekonstruktion på svensk kyrkoboksmaterial. Problem och möjligheter". *Meddelande från Ekonomisk-historiska institutionen,* Lunds universitet, Nr 59, Lund. See also Bengtsson, T and Lundh, C (1994), "Child and Infant Mortality in the Nordic Countries". *Annales de démographie historique.* Paper presented at an IUSSP conference in Montreal 1992.

23 Bengtsson, T and Ohlsson, R (1985) "The Standard of Living and Mortality Response in Different Ages", in *European Journal of Population,* vol 1(4), Bengtsson, T (1993), "Combined Time Series and Life Event Analysis" in Reher, D and Schofield, R (eds) (1993) *Old and New Methods in Historical Demography.* Oxford and Bengtsson, T (1993) *Economic, Social and Biological Determinants of Infant and Child Mortality in a Swedish Mining Parish 1757-1850. Combined Time Series and Life Event Analysis.* Paper presented at the IUSSP General Conference in Montreal. Lund.

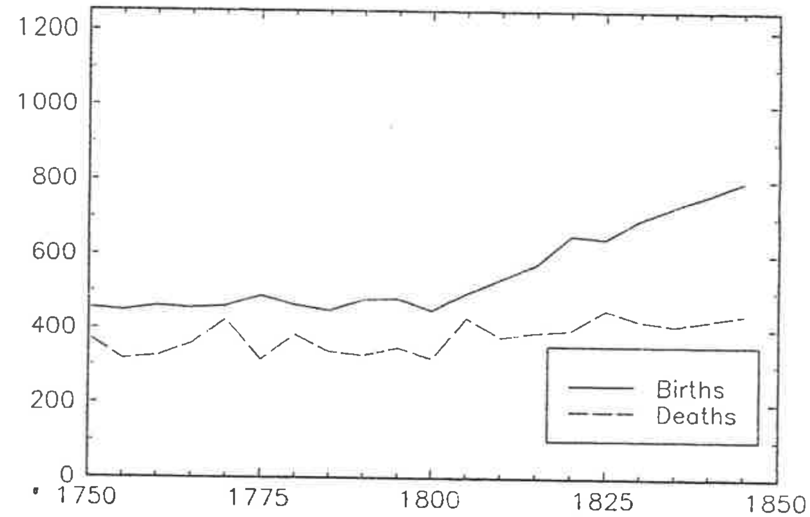

Figure 1. Births and deaths in Scania sample 1750-1850. Annual averages for quinquennia.

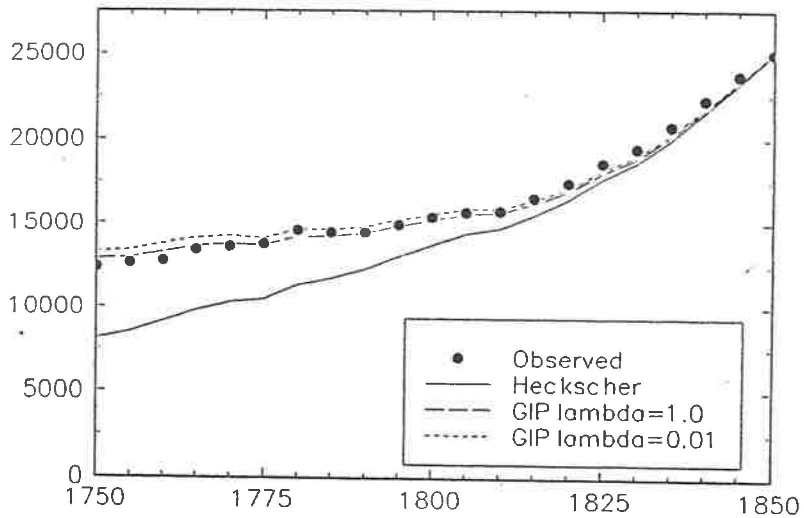

Figure 2. Scania sample population estimates 1750-1850. Alternative experiments. Population totals.

increase is particularly strong after the Napoleonic wars when births boomed, as shown in Figure 1, which shows 5 year totals of the number of births and deaths for the 29 parishes used for this period. There is also a net out-migration throughout the period.

The results of two GIP projections are shown in Figure 2, together with the observed population and a back projection using Heckscher's method, i.e. assuming closed population. The two GIP projections made full use of the information about the age structure in 1850. Thus, information on population in the ages 0 to 84, i.e. in 17 age groups, were used. The 18th age group has an open interval and has therefore been excluded. Thus persons in the highest age group with a closed interval were born almost at the start of the projection or, more precisely, in 1765-70, so that the number of time intervals is only slightly larger than the number of age groups. The two GIP experiments differ with respect to the importance of the migration term in the equation we are minimising. In one of the cases lambda is set to be equal to 1, which means that differences in migration-rates in subsequent five year periods were regarded as being as important as the differences in observed and estimated deaths and observed and estimated number of persons in the age-groups of the terminal census. In the other GIP projection lambda is set to 0.01, which means that migration changes were regarded as of almost no importance when projecting the population.

The results of the two GIP experiments differ only slightly and, comparing the estimates with the observed populations, we find that the projection with lambda equal to 1 is slightly better. Estimated total populations differ from the recorded totals by, on average, only a few percentage points. The difference is greatest before 1770, which may be a result of mismatch between the form of our survival table and actual survivorship in the early 1770s, when two subsequent harvest failures in combination with an epidemic outbreak caused numerous deaths. However, the inaccuracy of the projection is never worse than 3.8 per cent. With lambda equal to 0.01 the differences are somewhat greater than with lambda equal to 1. Compared with Heckscher's method, where the population is underestimated by 35.1 per cent in 1750, the results are very good.

In the GIP experiment with lambda equal to 1, the largest errors are observed for the group aged 75-79 years in the 1850 census, for deaths and migration in 1770-75, and for migration in 1785-90. That the errors are greatest for higher ages questions the quality of the data in the census, but those aged 75-79 in 1850 were born in the early 1770s. Thus, it would appear that the epidemic outbreak in that period creates problems for the estimation. It seems that the survival schedule used may not be appropriate for periods with epidemic outbreaks. A possible solution would be to change the survival table in periods when deaths exceed births. However, the size of the problem is not alarming since the discrepancies between the estimated and observed population are never worse than 3.8 per cent. Thus, it is obvious that the sample is large enough for GIP and that the data is good enough.

With lambda equal to 0.01, i.e. when fluctuations in migration are not considered important in calculating population size, the error pattern changes a great deal. Firstly, estimated deaths are closed to observed deaths. Secondly, errors in age structure are not observed at higher ages but in the group aged 10-14 years. Most fluctuations in migration are in the period around 1770, but also occur 1830-35. During the latter period deaths are high and the survival population age structure is peculiar in comparison with the one of 1850. But, as before, estimated populations are close to the observed ones.

The difference between estimated and observed age structure also indicates problems with age-heaping in the 1860 census, as the signs of the errors change for each successive age group. However, analysing subsequent censuses we find that age-heaping successively changed over time. Our conclusion is therefore that we don't have a problem with age heaping in our data. The changing sign in the difference between estimated and observed population age groups is likely to be the result of variations in births. Earlier studies have shown that cycles of 6-8 years are prevalent in the birth series.[24]

Estimates of Population from 1650 to 1760

The data for the GIP reconstruction from 1650 to 1760 are similar to the ones used above with respect to population size and length of time period, but the experiment differs in other respects. Firstly, we have inflated the series at the beginning of the period as the numbers of parishes with information about births and deaths increases over time.

24 Bengtsson, T (1984), "Harvest Fluctuations and Demographic Response, Southern Sweden 1751-1859", in Bengtsson, T, Fridlizius, G and Ohlsson, R (eds) (1984) *Pre-Industrial Population Change - The Mortality Decline and Short-Term Population Movements*. Lund.

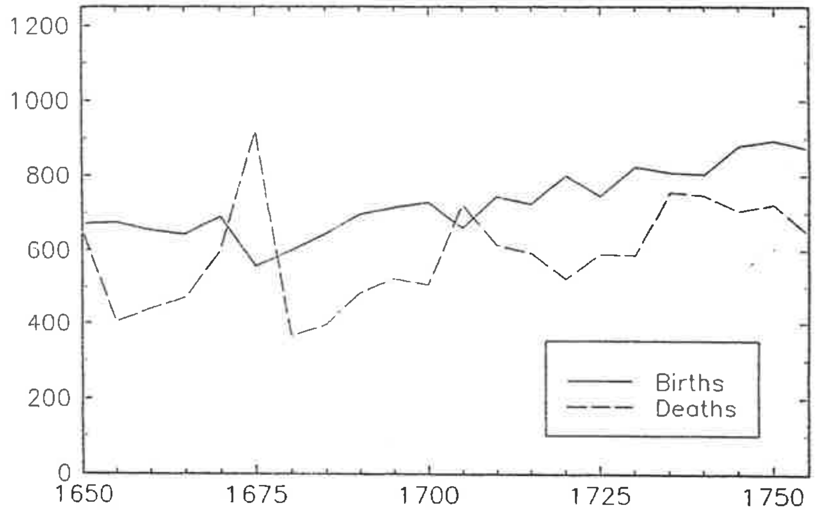

Figure 3. Births and deaths in Scania sample 1650-1760. Annual averages for quinquennia.

Secondly, there are more epidemics. Some of the parishes experienced a plague in 1654, an epidemic outbreak in combination with the war in 1677 and another plague in 1711 (Figure 3).[25] Unfortunately, the mortality peaks in the first two cases take place in the period when we have data for rather few parishes. There are only nine parishes in observation in 1651 and only eleven between 1671 and 1685. The number of parishes then increases substantially. The problem is similar to that for the reconstruction of the English population back to 1541, but at a different level - the sample is much smaller in our case.

In the GIP estimations for this period we will also make use of a target population in order to improve the quality of the estimation. We will use the taxation records for 1699 and 1700 which give information about the population aged 15 years and above. We are not using the summaries that were reported but the actual taxation records, which allows us to include poor persons, the disabled and other persons that were not always included in the summaries. Still, we have

25 For the plague in 1711, see Moseng, O G (1990), *Nordens siste pestepidemi: En punktstudie av Allerum 1710-1711*. Hovedoppgave i historie, Universitetet i Oslo.

reason to believe that not all persons are included and that the target population in 1700 is a slightly low one.

Experiment 1 used the basic GIP model: the census age-groups in 1760 and the quinquennial totals of deaths were used as targets for the estimation process and the migration was smoothed with lambda equal to 1. The population estimate for 1650 was 22,009, which seemed impossibly high compared with the scattered information that we have from taxation records around that time. The program's diagnostics indicated that the data were inconsistent with the model and its assumptions. The deaths total for 1675-1680 (the major epidemic period) was under-estimated by 18 per cent and this period also showed the biggest change in net-migration rates. Paradoxically, the 1675-1680 cohort in the terminal census, who were 80-4 years of age in 1760, was under-estimated by 20 per cent, suggesting that infant deaths were over-estimated even though not enough deaths were estimated for the period to match the observed number. This appeared to be strong evidence that the age-structure of mortality in this crisis did not match the shape of the model schedule.

Before adjusting the model, in Experiment 2, we used the target described above within the same model. This had a major impact on the estimated populations, bringing them down to plausible values for the seventeenth century, but did not resolve the major discrepancies arising from the 1675-80 epidemic. It also showed that the model over-estimated the 15-plus population in 1700 by 19 per cent.

A number of changes to the basic model were examined. The most promising results came from models where the migration smoothing was "loosened", with lambda set at .1, and where the shape of mortality was selected by the model for epidemic years. GIP uses a Brass relational model for mortality. This model has two parameters, one for the level of mortality and one for its shape. Because the historical population reconstruction problem already has too many parameters to be estimated, the basic GIP model only allows the level parameter to be changed for each time-period and fixes the shape parameter at unity. Thus each period's mortality can vary in level from the single input mortality schedule, but always retains its shape.

Experiment 8 allowed the model to estimate the mortality shape parameters for 1675-80 and 1705-1710, instead of fixing them at 1.0. For 1675-80, the estimate was 2.22, whereas for 1705-10 it was 1.10. The diagnostics now showed that the 80-4 year olds in the 1760 census were matched very well and that the errors in estimating the

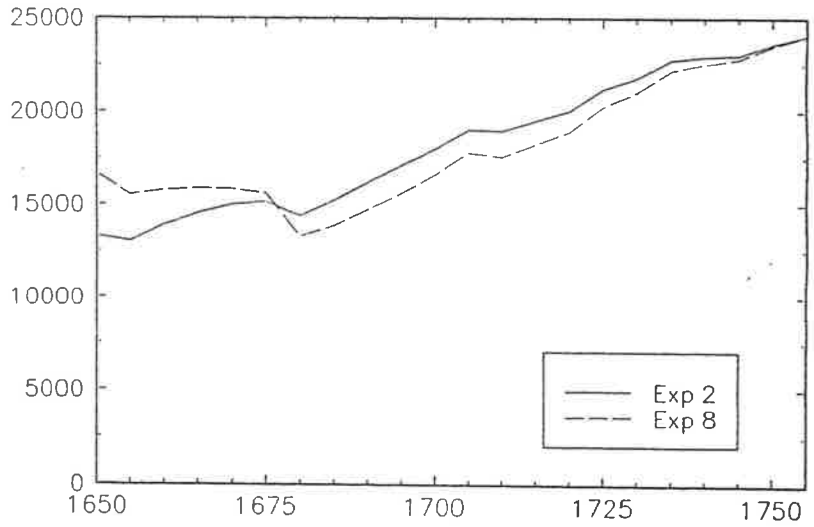

Figure 4. Scania sample population estimates 1650-1760. Alternative experiments. Population totals.

1675-80 deaths and the migration fluctuation in that period had all disappeared. The remaining fluctuations in migration were confined to the early years of the data - 1650 to 1665. The over-estimation of the adult population in 1700 was only 7.0 per cent. As we believe that the target in 1700 is slightly low we feel very confident with this result. The estimation of the population size back to 1675 also seems plausible (see Figure 4). In the period before 1675 it seems that the estimation is too high due to the problems discussed earlier.

Experiment 12 allowed a mortality shape parameter for 1670-5, a period when deaths do seem to be higher than the background trend. The parameter was .74, which suggests that the mortality in that period might have been more inclined towards infants, if we discount the possibility of age-overstatement in 1760, but it only reduced the discrepancy for the 90-4 year-olds to 5.6 per cent.

These results suggest that there was indeed a major shift from the normal pattern of mortality towards non-infant deaths in the earlier period, but that the later epidemic appeared to be the normal pattern, but at a more severe level. Experiment 9 dropped the later shape-

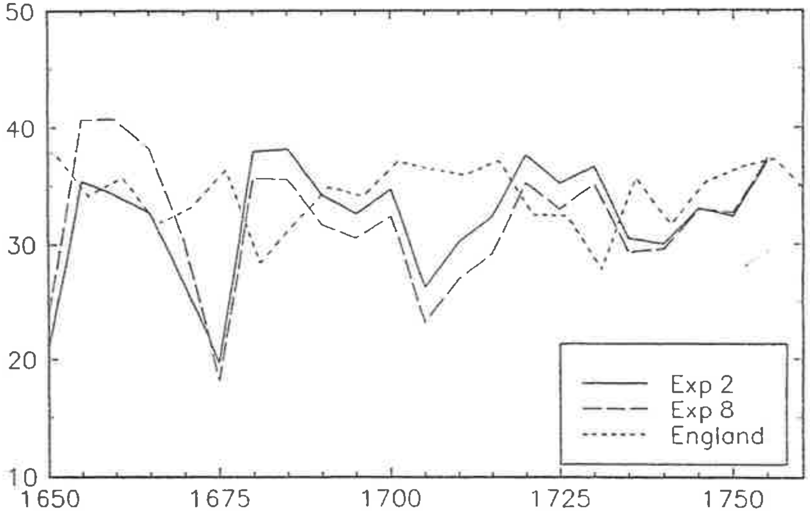

Figure 5. Estimated Life-Expectancy for Scania sample and England 1650-1760.

parameter, which caused only a slight shift in the earlier estimate to 2.45.

The 1675-80 epidemic presents a particular problem for the GIP model, as the cohorts above age 25 are being reconstructed backwards since the sizes of their birth-cohorts are unknown. For example, the deaths allocated to the 50-5 age-group in the 1680 census are subtracted to get the 45-9 group in 1675. The shape parameter in the Brass relational model acts in such a way
that adjustment to the mortality rates is an increasing function of age. This means that, once GIP decides to make the age-pattern of mortality "older", the inflation of the elderly cohorts in 1680 is quite marked and an unlikely ridge appears in the sequence of censuses, tracking backwards from late ages in 1675 to the middle-aged in 1650 when the data begins. Since we believe that these deaths should have been allocated to boost the middle years of adult life, we suspect that the early estimates of fertility may be too high and the estimates of mortality too low. The exact assessment of the magnitude of the problem must wait until our assumptions about the age-patterns of

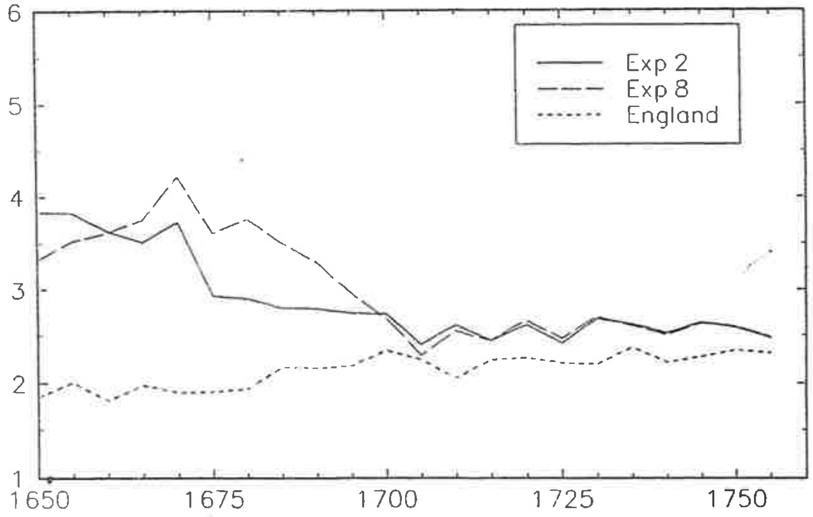

Figure 6. Estimated Gross Reproduction Rates for Scania sample and England 1650-1760.

Source: For England, Wrigley, E A and Schofield, R S (1981), *The Population History of England 1541-1871 - A Reconstruction*, London.

crisis mortality are confirmed and incorporated into the mortality model.

Comparing experiments 2 and 8 we have found that the last one comes closer to the population target in 1700 and that it also fits the changes in the mortality pattern in 1705-10 better. Life expectancies at birth estimated from these two experiments are shown in figure 5 together with the life expectancies for England. The curves are quite close after the great Nordic war, which is as expected. Before then the estimate for Scania are very different. It is not so much that the level is different, although it is on average lower for Scania during the 17th century, as it is that the values for Scania are changing much more from one period to the next. That is not to say that they are implausible. On the contrary, we expect the variation to be larger for a smaller population than a large one. Furthermore, and more

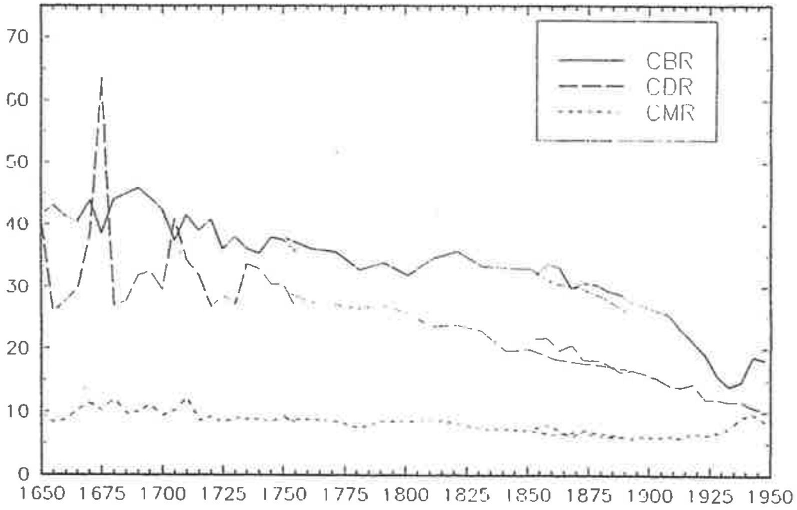

Figure 7. Crude Birth, Death and Marriage Rates for Scania sample, Scania and Sweden. Rates Per Thousand.

Source: For Sweden *Historisk Statistik för Sverige, del 1 Befolkning* (1969), 2:a uppl, Stockholm, for Scania 1750-70 and 1800-1900 Sundberg, G, *Emigrationsutredningen. Bilaga V.* (1910), Tab 31, Stockholm, and for Scania 1770-1800 Fridlizius, G, unpublised data for a sample of 80 Scanian parishes.

important, is that we know that Scania experienced both severe epidemics and wars during until the end of the Nordic war. The low values in the first two decades of the 18th century, for example, seem very plausible. The same is true for the 1675 and 1650 periods. Furthermore, the divergence between the two experiments is small after 1675.

When it comes to the estimates of gross reproduction rates, shown in figure 6, the curves are much smoother. The fertility is higher in Scania than in England during the entire period, but only slightly higher after 1721. The divergence between the two experiments is quite large for the 17th century but the results are, in general, consistent with our knowledge about this period.

Figure 7 shows the crude birth and death rates for Scania and Sweden using estimates from experiment 8. Given the uncertainties discussed earlier, it seems that Sweden experienced two periods of falling mortality and fertility, the first period starting around 1700. The beginning of the 19th century seems to be another period of change. The mortality level is increasing and the fertility level decreasing during the Napoleonic wars. Thereafter births are booming and mortality dropping. Thus the results are rather similar to England during this period. The marriage rate, which comes out as on independent result since marriage data is not used by GIP, is as expected. Our results must, however, be scrutinised further before we start to rewrite the Scanian population history. First, we need to model the epidemics in 1675 better. Then we must compare our results with other sources, however scattered.

In a first attempt to do the latter we have analysed the deaths in 1677 in the two parishes that experienced the mortality outbreak. Mortality started to increase in December 1676 and peaked in the first months of 1677. No information is given about cause of death or age at death. By using family reconstitution we were able to estimate the age of death within a 5 year age group, or at least whether the person who died was a child or an adult, for 80 per cent of the deaths. This showed that the mortality was much higher than normal for children and young adults and confirms our conclusion that the standard mortality schedule that we used is inappropriate for crisis mortality. In our further research we will make use of special schedules for this period and perhaps for other periods with extremely high mortality.

Summary

Our results have shown that is possible to make use of aggregated births and deaths totals from a rather small sample of parishes as a source for estimating population size in the period before our censuses begin. In our experiment we estimate the population of Scania for the period 1650 to 1760. The method used is Generalised Inverse Projection and the data used, besides births and deaths, are the population age structure at the end of the period, and mortality, fertility and migration schedules. The existence of a population target back in time makes it much more convincing to estimate populations that fit the series of births and deaths that are used as inputs. Here, we use information about the population aged 15 years and above in the taxation records in 1700. In our case we are only going back 110 years in time, which means that the information from the age-structure in 1760, covers almost the entire period. This also narrows the band of possible populations. Although some modifications and verifications are still to be made, we will argue that our results are conclusive, in particular for the period after the mortality crises in 1675. Our basic finding is that the mortality and fertility levels prior to the demographic transition are far from stable, as the theory of the demographic transition assumes. Thus Scania, the southernmost part of Sweden, seems to have a similar population development to England and France in this respect, although we must remember that this is a relatively small region.

APPENDIX

Population in 5-year age groups in 1760 and the migration, survival, and fertility schedules used in the GIP for 1650-1760.

AGE	POP 1760	POP 1850	P(MIGR)	P(SURV)	P(FERT)
0	3290	3419	0.0650	0.8053	0.0000
5	2654	2901	0.0450	0.8953	0.0000
10	2538	2599	0.0400	0.9630	0.0000
15	2099	2525	0.0650	0.9716	0.0250
20	2050	2018	0.2000	0.9628	0.1275
25	1901	2161	0.2800	0.9546	0.2525
30	1708	1731	0.2000	0.9500	0.2650
35	1676	1550	0.0750	0.9448	0.2000
40	1449	1270	0.0300	0.9380	0.1075
45	1274	1176	0.0000	0.9289	0.0225
50	1098	1002	0.0000	0.9152	0.0000
55	904	929	0.0000	0.8925	0.0000
60	796	650	0.0000	0.8551	0.0000
65	602	482	0.0000	0.7966	0.0000
70	425	332	0.0000	0.7085	0.0000
75	203	206	0.0000	0.5914	0.0000
80	113	70	0.0000	0.4592	0.0000
85	33		0.0000	0.3168	0.0000
90	10		0.0000	0.1868	0.0000

Sources: See text.

Lund Papers in Economic History

1990

1. Tommy Bengtsson & Christer Lundh, Life Events in a Peasant Society in Transition. Causes of Population Change in Scania, 1660-1860. */Population Economics/*
2. Tommy Bengtsson, Does Generation Size Affect Migration. Internal Migration in Sweden 1961-1988 by Age and Sex. */Population Economics/*
3. Lars Svensson, Arbete, löner och kön. En studie av befolkningsstruktur och relativa löner i kontorsyrken vid mitten av 1930-talet. */Education and the Labour Market/*
4. Lars Pålsson Syll, Adam Smiths fördelnings- och tillväxtteori. */General Issues/*
5. Anne Jerneck & Ingela Palmgren, Competitiveness and Consensus. Reflections on Swedish and Latin American Development Experiences in a Comparative Perspective. */Development Economics/*

1991

6. Göran Ahlström, London '51 - Paris '55 - London '62, International Industrial Exhibitions and Swedish Industry in the Middle of the 19th Century. */General Issues/*
7. Christer Lundh, The Crisis of the Swedish Model. Recent Trends in Collective Bargaining in Sweden */General Issues/*
8. Tommy Bengtsson & Christer Lundh, Evaluation of a Swedish Computer Program for Automatic Family Reconstitution. */Population Economics/*
9. Lars Pålsson Syll, Pre-klassisk fördelnings- och tillväxtteori. */General Issues/*
10. Jonas Olofsson, Rationaliseringsideologi, arbetsmarknadspolitik och utbildning. Rehn-Meidnermodellen i ny belysning. */Education and the Labour Market/*
11. Lars Pålsson Syll, Den svenska modellen - lön och samhällskontrakt. */General Issues/*
12. Anders Nilsson & Birgitta Svärd, The Quantitative Development of Vocational Education in Sweden 1950-1990. */Education and the Labour Market/*

1992

13. Anders Nilsson & Lars Pettersson, Education, Knowledge, and Economic Transformation. The Case of Swedish Agriculture 1800-1870. */Education and the Labour Market/*
14. Göran Ahlström, Technical Competence and Industrial Performance. Sweden in the 19th and early 20th Centuries. */General Issues/*
15. Tommy Bengtsson, Den demografiska utvecklingen i de nordiska länderna 1720-1914. */Population Economics/*
16. Lars Svensson, Ekonomisk strukturförändring och kvinnolöner. En diskussion om orsaker till variationerna i kvinnliga privattjänstemäns relativa löner 1955-1990. */General Issues/*
17. Lars Pålsson Syll, David Ricardos fördelnings- och tillväxtteori. */General Issues/*
18. Anders Nilsson & Lars Svensson, Familjeekonomi och utbildningsval. */Education and the Labour Market/*

1993

19. Lars Pettersson, Yrkesutbildning, ekonomisk omvandling och tillväxt i Sverige under efterkrigstiden. */Education and the Labour Market/*
20. Anders Nilsson & Birgitta Svärd, Skrivkunnighet på landsbygden i Skåne under tidigt 1800-tal. Källmaterial och preliminära resultat */Education and the Labour Market/*
21. Mats Didriksson, Arbetets lön. En studie av löneandelens utveckling inom svensk industri 1952-89. */General Issues/*
22. Jonas Olofsson & Rolf Ohlsson, Labour market policies in the interwar years: passive and active actions. */General Issues/*
23. Göran Ahlström, Industrial Research and Technical Proficiency. Swedish Industry in the Early 20th Century. */General Issues/*
24. Mauricio Rojas, Iransk, latinamerikansk och polsk företagarverksamhet i Malmöhus län 1989-1992. */General Issues/*
25. Neelambar Hatti & James Heimann, The State and Local Revenue Administration i South Indian History; Yelendur Kaditas and the Role of Village Accountants and Village Headmen. */Development Economics/*

26. Benny Carlson, Den enprocentiga revolutionen. Debatten om riksbankens ställning i samband med räntekuppen 1957. */General Issues/*
27. Ingela Schånberg, Den kvinnliga utbildningsexpansionen 1916-1950. Realskolestadiet. */Education and the Labour Market/*
28. Somboon Siriprachai, Rent-Seeking Activities - A Survey of Recent Issues. */Development Economics/*
29. Tommy Bengtsson & Christer Lundh, Name-Standardisation and Automatic Family Reconstitution. */Population Economics/*
30. T S Faran, Economic Development and Islamic Response. The Case of Iran, 1962-1992. */Development Economics/*
31. Christer Lundh, Giftermålsmönster i Sverige före det industriella genombrottet. */Population Economics/*
32. Tommy Bengtsson & Jim Oeppen, A Reconstruction of the Population of Scania, 1650-1760. */Population Economics/*