IFAU – INSTITUTE FOR LABOUR MARKET POLICY EVALUATION

Essays on inequality and education

Erika Ekström

DISSERTATION SERIES 2003:2

Presented at the Department of Economics, Uppsala University

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Industry, Employment and Communications, situated in Uppsala. IFAU's objective is to promote, support and carry out: evaluations of the effects of labour market policies, studies of the functioning of the labour market and evaluations of the labour market effects of measures within the educational system. Besides research, IFAU also works on: spreading knowledge about the activities of the institute through publications, seminars, courses, workshops and conferences; creating a library of Swedish evaluational studies; influencing the collection of data and making data easily available to researchers all over the country.

IFAU also provides funding for research projects within its areas of interest. There are two fixed dates for applications every year: April 1 and November 1. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The authority has a traditional board, consisting of a chairman, the Director-General and eight other members. The tasks of the board are, among other things, to make decisions about external grants and give its views on the activities at IFAU. A reference group including representatives for employers and employees as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

This doctoral dissertation was defended for the degree of Doctor in Philosophy at the Department of Economics, Uppsala University, December 2, 2003. The third essay of the thesis contains a revised version of research previously published by IFAU as Working Paper 2002:23. The fourth essay has been published by IFAU as Working Paper 2003:16.

ISSN 1651-4149

Doctoral dissertation presented to the Faculty of Social Sciences 2003

Abstract

EKSTRÖM, Erika, 2003, Essays on Inequality and Education; Department of Economics, Uppsala University, *Economic Studies* 76, 166 pp, ISBN 91-87268-83-3.

This thesis consists of an introduction and four self-contained essays.

Essay I examines the contribution of socio-economic variables to Namibian income inequality. I examine the extent to which total income inequality is due to within-group inequality or between-group inequality. Income inequality in Namibia ranks among the highest in the world. The within-group inequality seems to be the principal determinant of total inequality. Education is an important factor in determining degrees of between-group inequality.

Essay II examines the differences in earnings between males and females in manufacturing, services and the public sector in Namibia. The estimated earnings differences are decomposed into endowment and discrimination components. The results suggest that females are discriminated, but that females have a productivity advantage over the males, which reduces the gross wage differential. Comparing the OLS results with the results accounting for selection, the endowment component is not affected, whereas the discrimination component is reduced.

Essay III evaluates the 1991-reform adding a third year in Swedish upper secondary vocational education. One purpose with the additional year was to facilitate university enrolment for students from vocational paths. Reduced forms are applied to estimate the effect of a third year on three outcomes: years of upper secondary education, university enrolment and the rate of inactivity. The results suggest positive effects on university enrolments within six years for individuals with a three-year vocational education, and negative effects on activity.

Essay IV evaluates adult secondary education (ASE) in Sweden. ASE offers courses at the compulsory and upper secondary level and is aimed to give adults who lack these types of education. Controlling for pre-programme annual earnings, the estimates suggests that participating in adult secondary education significantly reduces the earnings of native-born males. No effects are found for native-born females, but the results indicate weakly significant positive effects for female immigrants.

To my mother and father

Table of contents

Acknowledgement

The intention of my studies in economics was to only read the first four courses at the A-level. As you can all see – you hold the result in your hands – I failed! I can't say it has been easy, and I would like to compare my studies to a winding path through a deep forest of trees, where the trees symbolise equations and empirical problems. Fortunately, the forest path has been lined with people giving me support, advice and encouragement.

First, I would like to thank Arne Bigsten at Göteborg University, advisor on my first two essays; Per-Anders Edin (advisor) and Peter Fredriksson (co-advisor) at Uppsala University, advisors on my last essays, for support and advice.

The first guidance was during my summer job at Statistics Sweden. Knowing of my interest in development economics Chris Denell gave me advice to contact Lena Åström. Without their help and assistance, my visit to Namibia would never have been possible. In Namibia my research took form. I gained an insight into the Namibian situation from discussions with Staffan Wahlström and the staff at the Central Statistics Office.

Moving onwards down the path, the forest began to be filled with trees. Luckily, there was an "Arne-cabin" with an open door. I only had to drop in, to discuss my problems.

Since I wanted to move home to Danderyd[1] I applied for a job at IUI – The Research Institute of Industrial Economics. In March 1997, I joined IUI as a research assistant in the project "Education and Training: New Job Skill Needs and the Low-Skilled" directed by Eugenia Kazamaki Ottersten and Erik Mellander. They introduced me to the educational area, for which I am most grateful. During my time at IUI I defended my licentiate thesis on Namibia at Göteborg University. Eugenia and Erik provided continuous encouragement, support and advice in my research. Thank you so much for that!

At the time I joined IFAU – Institute for Labour Market Policy Evaluation I decided to resume my PhD-studies, and I applied for the PhD programme at Uppsala University. P-A and Peter became my advisors. It is surprising that two so different people can be so very much alike.

[1] Several comments on where I live have been made, and to all of you who have made such comments I just want to say: I would never change with you.

Both have given me guidance on the remainder of the path. P-A pointed out the important details, for which I am most grateful. Peter continuously read my drafts and gave outstanding comments. His insight in the labour market field has fascinated me and I have benefited from his excellent guidance.

IUI and IFAU have both provided me with a stimulating and excellent research atmosphere and special thanks therefore go to Ulf Jakobsson (Director of IUI) and Susanne Ackum (Director of IFAU).

Other persons I would like to thank along the way are: my course mates Javiera Aguilar and Torun Österberg (neé Larsson) for making the PhD-courses more amusing; Eva-Lena Neth Johansson and Eva Holst for various kinds of issues during my PhD-studies; Åsa Murray for her answers about vocational education; my IFAU-roommate Daniela Fröberg for being the caring political scientist; administrative personnel at IFAU, especially Margareta Wicklander for excellent layout support.

My friends, especially my non-economist friends have made my "forest-life" brighter! There are, however, two economist friends who I would like to thank. At times of hopelessness, Eva Mörk has had words of comfort. I wish especially to thank Inger Johannisson for fruitful discussions and her continuous patience along the path. A huge thank you to you!

Finally, I have reached the edge of the forest and can see the light on the other side, and this is mainly thanks to my family, Inga and Gillis Ekström, and my sister Sophia, who have been beside me throughout the path. They have continuously given me support and encouragement so I could find my way out of the forest.

Financial support from the Foundation in memory of Lars Hierta, the Adlerbertska scholarship fund and the Institute for Labour Market Policy Evaluation are also gratefully acknowledged.

Uppsala, September 2003
Erika Ekström

The more you study, the more you learn;
The more you learn, the less you know.
So, why study?

Introduction*

A thesis in economics usually consists of self-contained essays, but they are often very closely related. This thesis contains four self-contained essays covering two subject fields: development economics and educational economics. Although the fields are different, all four essays assess the labour market from different angels. The first two essays on income distribution and wage discrimination focus on Namibia, whereas the two other essays focus on the labour market effects of the Swedish educational policies.[1]

Essay I analyses the income distribution among the Namibia households. Using individual wage data from Namibia, wage discrimination is analysed in Essay II. Essay III deals with the effects of a third year in upper secondary vocational education. The effects of participating in adult secondary education on earnings are examined in Essay IV.

To help the reader in assessing the finding in the first two essays, this introduction begins by presenting a short historical background of Namibia. Similarly, since educational systems differ across countries the reader is briefly introduced to the Swedish school system, before the details of the results on the educational essays are presented.

Historical background of Namibia[2]

In 1884 the Germans declared Deutsch-Südwest-Africa[3] as their protectorate. The territory was divided into the so-called Police Zone, and the North/north-east territory. The former, with good land for farming and rich in minerals, was primarily for the white European settlers, and the latter, called the "homelands", was where the black indigenous African population lived.

Labour was required for the European-owned farms, to construct a basic infrastructure, and to extract minerals that had been found in the area. For this, a Labour Contract System composed of three parts was created. First, the

* Comments from Per-Anders Edin, Peter Fredriksson and Eugenia Kazamaki Ottersten are gratefully acknowledged.

[1] The two essays in development economics were defended on June 10, 1998 as a licentiate thesis at the Department of Economics, School of Economics and Commercial Law, Göteborg University.

[2] The information in this section is mainly based on United Nations (1986).

[3] This was the name of Namibia at that time.

Germans made agreements with the northern population's chief/king stipulating that the male workers would work for them in the "Police Zone" and, in exchange, the chief/king would receive gifts from the returning migrant workers. Second, legislative measures were introduced, such as "pass laws" that required all Africans to carry a pass. Third, restrictions in movement were imposed on the African population within or outside the "Police Zone". By 1913, 90 per cent of the Africans' adult males within the "Police Zone" were wage earners (United Nations, 1986, p. 34).

The German period ended in 1915 when South African troops on behalf of Great Britain captured Deutsch-Südwest-Africa. In 1920 the League of Nations granted Great Britain a mandate to administer Deutsch-Südwest-Africa, which then was renamed South West Africa. However, Great Britain passed the mandate on to South Africa. Certain obligations were required by the mandate concerning the welfare of the domestic population, but in general, South Africa ignored these conditions. The Namibians' hope that the land would be returned to them by the government of South Africa was not fulfilled. In fact, oppression increased with the result that the land in central and southern Namibia, with economically viable farming, was monopolised by the whites in 1926.

After 1945, the United Nations attempted to replace the mandate with a trusteeship, but South Africa did not accept this. From 1948 Namibia was treated as South Africa's fifth province. South West African People's Organisation (SWAPO) was established in 1960 to fight for independence. The United Nations General Assembly withdrew South Africa's League of Nations mandate in 1966. This was, however, ignored by South Africa, and in 1971 the International Court of Justice declared that South Africa's occupation of Namibia was illegal. The United Nations Security Council Resolution 435[4] of 1978 formed the basis for all subsequent negotiations on Namibia, but it took another decade before South Africa accepted the resolution. Finally, Namibia was able to gain independence in March 1990.

Income inequality in Namibia

Eriksen & Moorsom's (1989) survey reveals that much has been written about Namibia, but little focus has been paid on income inequality issues. The obvi-

[4] The resolution called for free and fairs elections under the control and supervision of the United Nations for Namibia as one political entity.

ous reason for the lack of research in this field is the lack of data, but Namibia's late independence in March 1990 has also contributed. Many authors writing about Namibia suggest that the country suffers from a severely skewed income distribution, but go no further in analysing it. A study by van Ginneken (1986) reports a Gini coefficient of about 0.66, but this figure relates to the middle of the 1970s. More recent information suggests that Namibia is the country with the highest Gini coefficient among the 110 countries included in Milanovic & Yitzhaki (2002).

Essay I analyses the income distribution among the Namibian households using the 1993/1994 Namibia Household Income and Expenditure Survey (NHIES). This survey provided data to make a study of income distribution possible. Little is known about the sources of income inequality besides the main underlying cause, which is the income difference between population groups due to the earlier apartheid system. In addition, other aspects are important in explaining and examining the distribution of income. The aim of Essay I is to investigate the unequal incomes in terms of socio-economic variables. What is the contribution of different socio-economic characteristics to income inequality?

To measure income inequality the Gini coefficient is employed. The relative contributions of within- and between-group inequality to total income inequality for the different socio-economic variables are also analysed. The urban/rural differences and differences in the level of educational attainment are often argued to be the main explanations of income inequality in developing countries. Is this also the case in Namibia?

The finding suggests that Namibia is a country with a highly unequal distribution of income. However, dividing Namibia into two major regions it is primarily the Central/southern region that suffers from a skewed distribution of income. It is the within-group inequality that is the major component in total inequality. Urban/rural differences do not appear to be a main source of income inequality. Educational attainment is a source of inequality, however.

One possible reason for the more skewed distribution of income in the Central/southern region may be that this region is more heterogeneous. The region is composed of economic sectors with a broad mix of activities involving unskilled, semi-skilled and skilled workers.

Wage discrimination in Namibia

Essay II deals with a novel topic in the context of less developed countries, namely wage discrimination. Assessing wage discrimination in a less developed African country Knight & Sabot (1991), Appleton *et al*. (1995) and Skyt Nielsen (2000) have used similar methods as employed in this essay. Essay II focuses on whether there exists a gender bias in wages against women within the Namibian labour market. The aim is to examine whether the differences in earnings between males and females are due to endowment differences or to discrimination. The techniques due to Blinder (1973), Oaxaca (1973) and Oaxaca & Ransom (1994) are used. These techniques rely, among other things, on non selectivity bias in the ordinary least square (OLS) estimates. Therefore, if systematic differences are present between individuals who have reported positive wages and those who have not reported any wages, the OLS estimates will be biased and inconsistent and this property is then carried over to the components of the decomposition. To correct for selection bias, Heckman's (1979) two-stage estimation procedure is also employed.

In this essay any unexplained differences between men and women are attributed to discrimination. This is a potential problem, but for wait of a better word I will still refer to this unexplained difference as "discrimination". Recent literature has made some progress in terms of directly estimating the extent of discrimination between men and women. The seminal paper by Goldin & Rouse (2000), for instance, suggests a significant improvement for female musicians from having blind rather than open auditions.

This essay provides earnings differences between males and females in three sectors: manufacturing, services and the public sector. The aggregation of these sectors is, however, also considered. In addition, there is a partitioning into urban and rural areas.

There are four main findings: i) females are better endowed than males; ii) females are discriminated and the discrimination is largest in the urban areas; iii) female workers in the public sector are less discriminated than in the manufacturing and the service sector; and iv) the selection bias seems to explain a large part of the discrimination.

Part of the explanation of the fact that the females are better endowed is that they have higher levels of education. The possible explanation of that female is discriminated may rely of the past segregation in the labour market due to the fact that the Labour Contract System only contracted males. That the females are more discriminated in urban areas may be explained by that they are more

productive in the rural areas compared with the urban areas. Further possible explanations of finding ii) may be that the labour market in the rural areas, compared with the urban areas, are more homogeneous, i.e. more similar jobs. Findings iv) suggests there is need to account for the selection bias when decomposing earnings differences into components of endowment differences and discrimination, at least in the Namibian case.

The Swedish school system

From the autumn of the year when a child turns six the municipalities must offer a place in a preschool class. However, there is also an option to skip the preschool class and enter directly into the first grade of primary education. The preschool lasts for one year. When the child turns seven he/she begins to take nine-year compulsory education.

The nine-year compulsory education is followed by the upper secondary education. Upper secondary education is voluntary. But 98 per cent of those who left the compulsory school in 2002 continued (National Agency for Education, 2003). Today's upper secondary school offers two academic and fifteen vocational programmes. In addition to these programmes individualised programmes are provided. All are three-year programmes. A student has the right to begin a programme in upper secondary school until the year the student turns 20. Individuals aged 20 or older who do not have a three-year upper secondary education or the equivalent, have the right to attend adult secondary education (ASE).[5] ASE offers basic education (corresponds to the nine-year compulsory education) and non-compulsory education (corresponds to the regular upper secondary education). In addition, post-secondary training programmes are offered for individuals who would like to obtain skills in a new occupation. ASE also includes Swedish for immigrants. These courses provide immigrants with knowledge of the Swedish language, society and working life.

When the requirements from upper secondary education or ASE are fulfilled the individuals may continue to university and institutions of higher education. Here, the students are offered specific programmes or individual courses of study.

[5] Adult secondary education is one form of adult education. Other forms of adult education in Sweden are, for example, staff training, on-the-job-training, labour market training and advanced vocational training.

The value of a third year in upper secondary vocational education

Changes in the educational system are often said to be required in order to keep up with the changes in the labour market. Such an educational change was taken in 1991 when the two-year vocational programmes in upper secondary education were prolonged with an additional year (Government Bill 1990/91:85).[6]

The objective of the reform was to give students at these programmes a broader education including more theoretical subjects. This would facilitate university enrolment for students from the vocational programmes and reduce the exposure to unemployment.

Before the Swedish Government decided on the 1991-reform, a pilot scheme with three-year vocational programmes including more workplace training was carried out. The pilot scheme was conducted in the academic years 1988/89, 1989/90 and 1990/91. The pilot scheme assignment is used as the exogenous variation in educational attainment, since conditional on a set of local characteristics it is argued in this essay that the assignment of the pilot scheme is random. More precisely, the differences in educational attainment caused by living in a municipality with the pilot scheme or not in 1990 are studied.

An enormous amount of research has been devoted to examine the causal effects of education on earnings (for a review of recent studies see Card, 1999). The literature seems not to distinguish between different educational paths. However, a recent study by Gill & Leigh (2003) investigates the rate of returns from vocational programmes in community college. The study suggests that enrolment in a community college vocational programme increases earnings.

Essay III provides additional evidence to the literature on vocational programmes, though at the upper secondary level and on other outcomes than earnings. The aim of Essay III is to examine the value of an additional year in upper secondary vocational education on: i) years of upper secondary education (including adult secondary education), ii) university enrolment, and iii) the rate of inactivity, where inactivity is defined as individuals who are neither employed nor in higher education. Register data on school leavers who completed a two-year upper secondary vocational education in 1992 and school leavers who

[6] Today the Swedish upper secondary school is again faced with new suggestions of changes (SOU 2002:120).

completed a three-year upper secondary vocational education in 1993 are used. Reduced form equations are estimated for each year of the period studied (1993–98). To estimate the causal effect of the additional year at upper secondary vocational education the Wald estimate (Wald, 1940) is used.

In brief, the effect is positive on all three accounts. Living in a pilot municipality: increases the probability of a third year at the upper secondary level; increases university enrolment within six years; and raises inactivity.

The first result is unsurprising, since the three-year vocational programmes only were available in the pilot municipalities. The effect is constant over time, suggesting that individuals with a two-year vocational education added a third year of education, through other educational paths, to the same extent regardless of pilot status. It seems that the reform was successful to increase university enrolment from vocational programmes, but not to reduce the exposure to unemployment. One possible reason for the less favourable outcome with respect to the exposure is the increase in unemployment rate among young people between 1992 and 1993. The school leavers in 1992, i.e. those having a two-year vocational education, may have managed to establish themselves on the labour market before it declined.

Earnings effects of participating in adult secondary education

Essay IV deals with a topic that we know little about; namely the labour market effects of participating in adult secondary education (ASE). ASE offers education at the compulsory and upper secondary level for individuals who do not have this kind of education or its equivalent. But ASE also offers post-secondary training programmes. ASE is open for individuals aged 20 and above, but priorities are given to adults with the shortest education.

It is noteworthy that there is only one study dealing with this issue (Alm Stenflo, 2000), given that ASE courses have been provided in Sweden since 1968. She found positive effects on both earnings and employment compared with non-participants. A few recent studies have, however, emerged because of the large investment within ASE in the form of the "Adult Education Initiative" (AEI)[7]: Westerlund (2000) and Axelsson & Westerlund (2001) who evaluate

[7] The AEI was introduced in July 1997 and continued until December 2002. The AEI was aimed at unemployed people aged 25–55 who did not have three-year upper secondary education.

the AEI on unemployment, and Stenberg (2003) who evaluates the short-run earnings effects of the AEI. Stenberg (2003) found negative short-run effects compared with participants in labour market training.

The motives for participating in ASE may vary between individuals, but the motives probably include increasing the individuals' possibilities to change their situation in the labour market, e.g. through a change of occupation. Such changes, in combination with a higher level of education, may in the long-run also lead to changes in income levels. The aim of Essay IV is to estimate the long-run effects of ASE on earnings.

Looking at the participation pattern, females have always participated in ASE to a greater extent than males (Statistics Sweden, Statistical Yearbook for Sweden). For that reason this essay will have a gender focus in order to separate male and female earnings effects of participating in ASE. There is also a division of the participants into native-born and immigrants, because native-born and immigrants may have different motives for entering ASE.

I have access to unique longitudinal register data. These data provide me with information on pre-programme earnings. Thus, I simply apply ordinary least square to estimate the earnings effects. This relies on the assumption that pre-programme earnings picks-up any (unobserved) differences between participants and non-participants. In addition, the non-participants are re-weighed so that they have the same distribution of characteristics as the participants.

The main findings are the following: There are small, significantly negative effects for native-born males, but no effects for native-born females. For the immigrant females, there is a positive, and borderline significant effect. These results suggest that it is important to estimate effects for males and females separately, because the participation in ASE affects the earnings of the males and the females differently.

It seems that the long-run earnings effects of ASE are rather dismal, at least for the native-born population. This may suggest that the value of larger-scale interventions such as the Adult Education Initiative (AEI) is limited. Thus, it seems to me that the long-run effects of the AEI are an important topic for future research.

References

Alm Stenflo, G (2000) Inkomst- och sysselsättningseffekter av kommunal vuxenutbildning, Temarapport 2000:1, Prognosinstitutet, Statistics Sweden, Stockholm.

Appleton *et al.* (1995), "Gender Differences in the Returns to Schooling in Three African Countries", Centre for the Study of African Economies, mimeo, Oxford.

Axelsson, R & O Westerlund (2001) "Inflöde och varaktighet som arbetssökande efter avslutad utbildning – Deltagare i Kunskapslyftet och arbetsmarknadsutbildning höstterminen 1997", Umeå Economic Studies no. 566, Umeå University, Umeå.

Blinder, A S (1973) "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, vol. 8, no. 4. pp. 436–455.

Card, D (1999) "The Causal Effect of Education on Earnings", in Ashenfelter O & D Card (eds.), *Handbook of Labour Economics*, pp. 1801–1863.

Eriksen, T L & R Moorsom (1989) "The Political Economy of Namibia, An annotated critical bibliography", The Scandinavian Institute of African Studies, Second Edition, Norwegian Foreign Policy Studies no. 69.

Gill, A M & D E Leigh (2003), "Do the Returns to Community Colleges Differ between Academic and Vocational Programs?", *Journal of Human Resources*, vol. 38, no. 1, pp. 134–155.

Goldin C & C Rouse (2000) "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians", *American Economic Review*, vol. 9, no. 4, pp. 715–741.

Government Bill (1990/91:85) *Växa med kunskaper – om gymnasieskolan och vuxenutbildningen*.

Heckman, J J (1979) “Sample selection bias as a specification error”, *Econometrica*, vol. 47, no. 1, pp. 153–161.

Knight J B & R H Sabot (1991) “Labor market Discrimination in a Poor Urban Economy”, in Birdsall N & R Sabot (eds.) *Unfair Advantage, Labor Market Discrimination in Developing Countries*, World Bank regional and sectoral studies, Washington D.C.

Milanovic, B & S Yitzhaki (2002) ”Decomposing world income distribution: Does the world have a middle class?”, *Review of Income and Wealth*, vol. 48, no. 2, pp. 155–178.

National Agency for Education (2003) *Statistics for child care, school and adult education 2003:2, Children, Staff, Pupils and Teachers*, Official Statistics of Sweden, Stockholm (in Swedish).

Oaxaca, R (1973) “Male-female wage differentials in urban labor markets”, *International Economic Review*, vol. 14, no. 3, pp. 693–709.

Oaxaca, R L & M R Ransom (1994) “On discrimination and the decomposition of wage differentials”, *Journal of Econometrics*, vol. 61, no. 1, pp. 5–21.

Skyt Nielsen, H (2000) “Wage discrimination in Zamibia: an extension of the Oaxaca-Blinder decomposition”, *Applied Economics Letters*, no. 7, pp. 405–408.

SOU 2002:120 *Åtta vägar till kunskap – En ny struktur för gymnasieskolan,* Slutbetänkande från Gymnasiekommittén 2000.

Statistics Sweden, *Statistical Yearbook for Sweden*, various issues.

Stenberg, A (2003) *An evaluation of the adult education initiative relative labor market training*, dissertation thesis, Umeå Economic Studies no. 609, Department of Economics, Umeå University.

United Nation (1986), *Namibia: Perspectives for National Reconstruction and Development*, United Nations Institute for Namibia, Lusaka.

van Ginneken, W (1986) “Incomes and Wages in Namibia”, International Labour Office, Lusaka, Zambia, World Employment Programme, Southern African Team for Employment Promotions (SATEP), Working Paper.

Wald A (1940) “The fitting of straight lines if both variables are subject to error, *The Annals of Mathematical Statistic*, vol. 11, no. 3, pp. 284–300.

Westerlund, O (2000) ”Inskrivningstid vid arbetsförmedlingen efter avslutade studier – Deltagare i Kunskapslyftet och arbetsmarknadsutbildning 1997”, Umeå Economic Studies no. 526, Umeå University, Umeå.

Essay I: Income inequality in Namibia*

1 Introduction

Reducing income inequality is one of the four goals of Namibia's national development policy, spelled out after the independence in March 1990. The country's skewed income distribution has been a problem for some time, but the lack of comprehensive data on income before 1993 made the problem difficult to discern until recently (National Development Plan 1, 1995). In this paper I carry out an empirical analysis on the household income distribution using the 1993/1994 Namibia Household Income and Expenditure Survey (NHIES). Differences between the population groups, an aspect more relevant in the past than it is today, may still exist in Namibia, but additional factors are important in explaining and examining the distribution of income. Two important aspects that are widely discussed in studies on income distribution in developing countries are the urban/rural differences and differences in the level of educational attainment. I investigate whether these two aspects are important in the case of Namibia.

The purpose of this paper is to examine the distribution of income among the households in Namibia. The main concern is to give answers to the question of what kinds of sources that contributes to income inequality in Namibia. Various socio-economic variables relating to the head of the household are examined, for example, age, urban/rural residence and the level of educational attainment. To measure the distribution of income the Gini coefficient is used and the income shares of the bottom 20 per cent of the population and the top 10 per cent of the population. To investigate to which extent total income inequality is due to within-group inequality or between-group inequality both Theil's (1967) entropy index T and Theil's second measure L are used. By means of these measures I find that the distribution of income in Namibia is very skewed. The skewness is more pronounced in the Central/southern region

* I am grateful for valuable comments from Arne Bigsten, Eugenia Kazamaki Ottersten, Erik Mellander, Åsa Stenman and seminar participants at the Department of Economics, School of Economics and Commercial Law, Göteborg University. The usual disclaimer applies.

than in the North/north-east region. The between-educational inequality seems to be the most important determinant of the skewed income distribution. Urban/rural residence and main source of income are also important variables in determining degrees of between-group inequality, particularly in the North/north-east region. Otherwise the within-group inequality seems to be the principal determinant of total inequality.

The remainder of the paper is outlined as follows. A socio-economic framework is first presented in Section 2. Section 3 reviews previous studies related to income distribution in Namibia. In Section 4, I discuss how to measure and decompose income inequality. The data and the income concept used are presented in Section 5. The results, focusing on the different socio-economic variables, are reported in Section 6. Section 7 contains a summary and conclusions.

2 Socio-economic framework

2.1 Economic overview

Namibia is one of the richest countries in Africa in terms of natural resources. It possesses diamonds, uranium, copper, lead, zinc and other minerals; rich marine resources; and a large stock of cattle and karakul sheep. The economy is small and mixed.[1] Being heavily dependent on exports of primary products (diamonds, uranium, base metals, cattle, karakul pelts and white fish), the Namibian economy is highly vulnerable to external shock such as fluctuations in world market prices.

Economic growth remained slow for most of the 1980s, due to unfavourable development of world market prices for karakul and mineral exports; constraints on investments; the poor performance of the South African economy; and the severe drought from 1980–84. The performance of the economy has improved since Namibia gained independence in 1990 showing only short interim of negative growth in 1993, about two per cent, because of drought (CSO, 1996b). The government and the primary sectors are still the main sources of economic growth (National Development Plan 1, 1995). *Figure 1*

[1] Mixed economy means that all the following forms of ownership are present: public, private, joint public-private, co-operative, co-ownership and small-scale family.

shows the percentage contribution to gross domestic product (GDP) by primary, secondary and tertiary industries during 1980–95.[2]

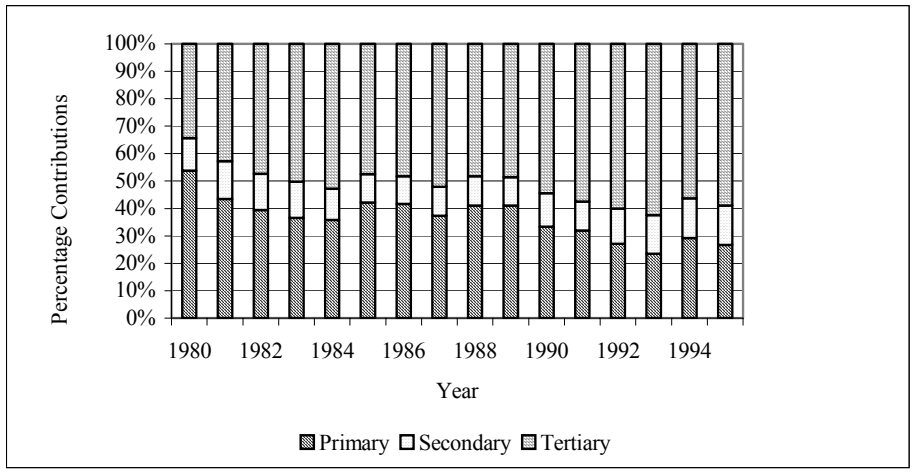


Figure 1 Gross Domestic Product by Activity in Current Prices, 1980–95

Source: CSO (1996b).

The relative contribution of the primary industries has decreased markedly from about 54 per cent in 1980 to about 27 per cent in 1995. The major part of this decrease is due to an unfavourable development of relative prices of primary sector output. In volume terms the share of the primary sector decreased only by nine percentage points, from 45 per cent in 1980 to 34 per cent in 1995. The droughts in the years 1980–84, 1992 and 1995, which affected agricultural production adversely, have contributed, but the decline is mainly due to the falling relative importance of the mining industry. Nevertheless, mining is still the key sector of the economy in terms of export earnings and government revenue. The secondary sector's share of GDP has been fairly constant over the period except, after 1991, when the share slightly increased to around 14 per cent. In general, manufacturing expansion has been held back by competition from South African firms (United Nations Industrial Development

[2] Primary industry includes agriculture, fishing, mining and quarrying; secondary industry includes manufacturing, electricity, water and construction; tertiary industry includes service, wholesale and retail trade.

Organization, 1990), but there was a slight increase after 1991, the result of an expansion in the fish processing industry. The tertiary sector has increased over the period, and it has been the most important sector since 1982. The main reason for the increase is the expansion in government services.

2.2 Social indicators

It appears that health, education and living conditions in Namibia have not improved at the same rate as the economy has grown in the 1990s. In 1994 the GDP per capita was US$ 1 970, which is high for a country in sub-Saharan Africa. For example, in Mozambique GDP per capita was US$ 90, in Kenya US$ 250 and in Zimbabwe US$ 500. In fact, only Botswana and South Africa have higher GDP per capita, US$ 2 800 and US$ 3 040, respectively (World Bank, 1996). However, regarding social conditions, the Human Development Report (1996) ranks Namibia 116 out of 174 countries according to the Human Development Index (HDI), see *Table 1*.[3]

Table 1 HDI rank and real GDP per capita rank for developing countries 1993

Country	HDI	HDI rank	Real GDP per capita (PPP$) rank
Botswana	0.741	71	60
South Africa	0.649	100	93
Namibia	0.573	116	79
Zimbabwe	0.534	124	120
Kenya	0.473	128	136
Zambia	0.411	136	144

Note: PPP is Purchasing Power Parity.

Source: UNDP (1996).

The HDI indicates that the average level of human development in Namibia is low by world standards. When compared with other sub-Saharan African countries, Namibia is not so poorly developed, but there is a huge difference between the ranking according to the HDI index and according to real GDP per capita.

[3] The closer the HDI is to one (1), the better is the country's human capabilities, i.e. a long and healthy life, knowledge and a decent economic standard.

During colonialism the educational system was organised along racial lines, and provided very unequal access to schooling. While education was compulsory and free of charge for the whites, the Africans were not required to go to school and if; nevertheless, they choose to do so they had to pay for their education. This educational inequality is still present in Namibia. However, the government has spent about 10 per cent of the GDP and about 27 per cent of government expenditures on education since independence.[4] A comparison of school attendance figures from the 1991 Census and the 1993/1994 NHIES is presented in *Table 2*.

Table 2 School attendance for population 6 years and older, per cent

School attendance	Urban		Rural		Total	
	1991 CENSUS	1993/1994 NHIES	1991 CENSUS	1993/1994 NHIES	1991 CENSUS	1993/1994 NHIES
Never attended	11	7	27	19	23	16
Still at school	31	32	40	45	37	41
Left school	58	60	33	36	40	43

Notes: Sample weights are used for 1993/1994 NHIES. Column two does not sum to 100 per cent due to a non-respond item.

Source: CSO (1995a) and CSO (1996a).

Table 2 suggests that the percentage for the category never attended has declined. However, one has to keep in mind the differences in the definition of the population.[5] This category has decreased from 23 per cent to 16 per cent, with the largest decrease being seen in the urban areas. How is the educational attainment distributed among females and males and age groups? *Table 3* shows that about one quarter of the population 6 years and older has no formal education.

[4] Own calculations using Tables F4 and B1, CSO (1996b).

[5] Recall that the 1993/1994 NHIES is based on private households, while the 1991 Census includes the institutional population.

Table 3 Educational attainment for population 6 years and older

Level of educational attainment	Females		Males		Total	
	Number	%	Number	%	Number	%
No formal education	144 520	24.0	129 537	24.8	274 057	24.4
Primary education	256 008	42.5	224 743	43.0	480 751	42.7
Secondary education	180 385	30.0	145 521	27.8	325 906	29.0
Tertiary education	12 399	2.1	12 808	2.4	25 207	2.2
Non-applicable	8 822	1.5	10 460	2.0	19 282	1.7
Namibia	**602 134**	**100.0**	**523 068**	**100.0**	**112 5202**	**100.0**

Notes: Sample weights are used. No formal education: No school experience or currently in Grade 1; Primary education: Grade 1–6; Secondary education: Grade 7–12; Tertiary education: Diplomas or certificate issued by a University or Technical Institutes, university courses leading to first degrees 1 to 4 years, post-graduate (diplomas, master degree, doctor degree) and teacher training.

One major finding in the table, contrary to expectations, is that females do not appear to be less educated than males. Note that in secondary education the females even outnumber the males. In tertiary education the number of males slightly exceeds that of females.[6] The same pattern of educational attainment is found when looking at urban and rural areas separately. In addition, disaggregating on regional level this pattern still holds for most of the thirteen regions. The number of females is higher in the tertiary education in Khomas, Ohangwena, Omusati and Oshana.

It is rather unusual in the African context that females have higher school enrolment than males. The fact that it is so is because males were often contracted for wage work or fought in the wars. Furthermore, since the chances recently have increased for females to receive a job in the areas of health and education – areas which require schooling – enrolment of females is increased. Looking at school enrolments in primary education as a percentage of age group, Namibia's female enrolment rate is the highest of the sub-Saharan countries (World Bank, 1997). In the case of secondary education South Africa has the highest female enrolments, 84 per cent compared with Namibia's 61 per cent.

[6] The findings are comparable with other official statistics in Namibia. See CSO (1995c), for example.

2.3 The population and the labour force

The population in Namibia is about 1.5 million.[7] The average annual population growth has been estimated to approximately three per cent for the period 1985–93 (Sparks, 1996). The population density is only 1.8 persons per square kilometre.[8] However, the density varies across the country at a considerable extent. The Oshana region in the North has a population density of about 31 persons, while in the Karas region in the South the population density is about 0.3 persons.[9] Since the country is arid, most of the population lives on the plateau. The majority of the population is black, 86 per cent, while coloured and white each represents seven per cent of the population (Murray, 1993/94). *Table 4* presents the distribution of households and individuals by region and gender (see Appendix A for a map over Namibia and its regions).

[7] According to the 1991 Population and Housing Census (CSO, 1995a) the institutional population counts to about 100 000 individuals. In the 1993/94 NHIES the population is estimated to be about 1.4 million, since the institutional population is excluded.

[8] Own calculations using the estimated population from the 1993/94 NHIES including the institutional population from the 1991 Census based on Namibia's official area of 824 292 square kilometres (Sparks, 1996).

[9] Own calculations using the estimated population from the 1993/94 NHIES based on Oshana's and Karas' official area of 5 180 and 162 384 square kilometres, respectively (CSO, 1995a).

Table 4 Households and individuals by region and gender

Region/ urban and rural areas	House-holds		Individuals Females		Males		Total	
	Number	%	Number	%	Number	%	Number	%
Erongo	16 611	6.8	36 787	5.0	37 607	5.7	74 395	5.4
Hardap	12 521	5.1	28 468	3.9	25 738	3.9	54 206	3.9
Karas	11 545	4.7	26 750	3.7	27 364	4.2	54 114	3.9
Khomas	34 101	13.9	78 393	10.7	83 360	12.7	161 754	11.6
Kunene	10 398	4.2	31 466	4.3	27 562	4.2	59 029	4.2
Omaheke	9 157	3.7	23 658	3.2	23 443	3.6	47 101	3.4
Otjozondjupa	22 827	9.3	49 563	6.8	50 875	7.7	100 438	7.2
Central/southern region	*117 160*	*47.9*	*275 085*	*37.6*	*275 949*	*42.0*	*551 037*	*39.7*
Caprivi	168 84	6.9	50 001	6.8	41 433	6.3	91 434	6.6
Ohangwena	25 574	10.4	106 632	14.6	84 226	12.8	190 858	13.7
Okavango	20 394	8.3	65 376	8.9	59 657	9.1	125 033	9.0
Omusati	21 822	8.9	86 081	11.8	66 949	10.2	153 030	11.0
Oshana	24 198	9.9	87 781	12.0	73 709	11.2	161 491	11.6
Oshikoto	18 795	7.7	60 604	8.3	55 531	8.4	116 134	8.4
North/north-east region	*127 667*	*52.1*	*456 475*	*62.4*	*381 505*	*58.0*	*837 980*	*60.3*
Urban	82 864	33.8	204 514	28.0	196 812	29.9	401 325	28.9
Rural	161 962	66.2	527 048	72.0	460 643	70.1	987 691	71.1
Namibia	**244 827**	**100**	**731 562**	**100**	**657 454**	**100**	**1 389 017**	**100**

Note: Sample weights are used.

For the household population the highest share, 14 per cent, is found in the Khomas region. The reason for the small population in Hardap and Karas is the aridity of the land, which severely limits development and investment. For Erongo and Omaheke the low density is mainly due to the fact that these regions are covered by desert, the Namib and Kalahari, respectively. The North/north-eastern regions have a larger share of the household population than the Central/southern regions.

Regarding the distribution of males and females by region, males are slightly over-represented in Khomas, but in the other Central/southern regions the distribution of the males and females is fairly equal. In the North/north-east regions, however, females are over-represented, due to earlier male labour migration to the Central/southern regions. Finally, more than 70 per cent of the population live in the rural areas.

How is the Namibian population distributed with respect to age? This is shown in *Table 5*.

Table 5 Age structure by gender

Age Group	Females		Males		Total	
	Number	%	Number	%	Number	%
0–14	300 719	41.1	292 634	44.5	593 353	42.7
15–24	156 204	21.4	138 576	21.1	294 779	21.2
25–64	235 354	32.2	194 866	29.6	430 220	31.0
65+	36 299	5.0	27 874	4.2	64 173	4.6
Non-applicable	2 986	0.4	3 505	0.5	6 491	0.5
Namibia	**731 562**	**100**	**657 454**	**100**	**1 389 017**	**100**

Note: Sample weights are used.

The young population dominates; about 65 per cent are under 25 years of age. The average age for females, males and the total population are 21.8, 23.1, and 22.5, respectively. The working population, defined as those aged 15–64, represents over 50 per cent of the population. Only about 5 per cent of the population are 65 years of age or older. The female population exceeds the male one by about 11 per cent, and the male/female ratio shows a deficit of males in all age groups. This under-representation of males was likely caused by the civil war prior to the independence or male migration to other countries. Comparing urban and rural areas with the total, the age structure is somewhat different. The main differences are found in the age groups 0–14 and 25–64. In the urban areas about 34 per cent are in the 0–14 age group and about 41 per cent are in the 25–64 age group. In rural areas the corresponding figures are about 46 and 27 per cent, respectively.

The labour market activities may affect the income distribution and therefore it is important to reveal the labour markets activities of the population. The population's activities are presented in *Table 6*. See Appendix B for definitions of the different activities.

Table 6 Activity status of the population

Activity status	Females		Males		Total	
	Number	%	Number	%	Number	%
Children under 15 years	300 822	41.1	292 724	44.5	593 545	42.7
Adults 15 years and over	430 519	58.8	364 731	55.5	795 250	57.3
Non-applicable	221	0	0	0	221	0
Total population	*731 562*	*100*	*657 454*	*100*	*1 389 017*	*100*
Economically inactive	220 063	51.1	133 740	36.7	353 802	44.5
Economically active	207 299	48.2	227 379	62.3	434 678	54.7
Activity not stated	3 157	0.7	3 612	1.0	6 770	0.9
Adults	*430 519*	*100*	*364 731*	*100*	*795 250*	*100*
Students	68 376	31.1	69 843	52.2	138 219	39.1
Homemakers	111 784	50.8	31 670	23.7	143 454	40.5
Income recipient	392	0.2	495	0.4	887	0.3
Disabled, old, retired, others	38 992	17.7	31 465	23.5	70 456	19.9
Non-applicable	520	0.2	267	0.2	786	0.2
Economically inactive	*220 063*	*100*	*133 740*	*100*	*353 802*	*100*
Fully employed	76 369	36.8	95 782	42.1	172 151	39.6
Under-employed	87 178	42.1	90 951	40.0	178 129	41.0
Unemployed	43 752	21.1	40 646	17.9	84 398	19.4
Economically active	*207 299*	*100*	*227 379*	*100*	*434 678*	*100*
Paid employee	70 993	43.3	121 799	65.2	192 792	55.0
Employer	713	0.4	2 356	1.3	3 069	0.9
Own account worker	37 156	22.7	32 133	17.2	69 289	19.8
Unpaid family worker	51 500	31.5	27 087	14.5	78 587	22.4
Non-applicable	3 186	1.9	3 358	1.8	6 544	1.9
Employed	*163 547*	*100*	*186 733*	*100*	*350 280*	*100*

Note: Sample weights are used.

About 55 per cent of the adults aged 15 years and over are economically active. Students and homemakers are the two largest groups in the economically inactive population. In the student category the females and males are rather equal in numbers. The economically active population, of which the males account for a larger share than the females, consists of the fully employed, the

under-employed and the unemployed. Over 50 per cent of the employed (fully employed and under-employed) are paid employees. This high share of paid employees is not surprising, since the people entered wage employment through the Labour Contract System. Another factor that also has contributed to the growth of wage employment is the policy of promoting export-oriented commercial agriculture. The production of food crops was disregarded by policy-makers and in turn the income for the subsistence farmers in the north decreased. Thus, the farmers have to depend on wage employment in the modern sector as a source of income.

Paukert & Robinson (1992) compare wage employment in Africa, Asia, Latin America and the Caribbean with that of three industrialised countries and they show that, with some exceptions, wage employment is much less significantly in Africa than in industrialised market economies. Namibia's percentage of the economically active population in wage employment is about 44 per cent which is comparable to Botswana's 41 per cent 1981 and Zambia's 43 per cent 1980. Paukert & Robinson (1992) also show that participation of females in wage employment is generally higher in Asia and Latin America than in Africa, and that the participation rate is still higher for the industrialised market economies. The Namibian female participation rate, about 37 per cent of total paid employees, is comparable with the Asian participation rate and thus high by African standards.

Employers and own account workers are both self-employed, where the latter do not have any paid employees, but can have unpaid family workers. The breakdown into employers and own account workers may be done according to differences in the type of jobs. For example, employers are more engaged in trade activities than are own account workers. Employers constitute about one per cent of the employed population, own account workers about 20 per cent, and unpaid family workers about 22 per cent. Females tend to dominate in the two latter categories, because they have traditionally taken care of the farming business.

The statistics of employment are generally not so reliable in less developed countries, and for this reason it is difficult to measure changes in the employment status. Still, unemployment in Namibia has certainly risen since independence, due to the return of demobilised combatants, and repatriation of exiles (World Bank, 1992). From the 1993/1994 NHIES I know that the males have higher employment participation rates, in all age groups, than the females, which is not surprising, because during the colonial period the females were

excluded from many types of economic activities and thus from the labour market. The rate of unemployment is generally higher for females than males. When under-employment is added to unemployment, the gender gap widens. This is so both at the national level and between the age groups, because the females tend to work more in own businesses that relate to agriculture activities, which is insufficient to keep them fully employed.

3 Previous studies of income distribution in Namibia

Some quantitative studies on income distribution have been made in Namibia for some of the thirteen regions that Namibia constitutes of. They are reviewed in this section.[10]

In the mid 1970s the Namibian population was grouped into about 245 000 households, where the Gini coefficient was estimated to 0.65, 0.67 or 0.69 depending on which household income measure is used (van Ginneken, 1986). The poorest 40 per cent of the households received 5.9 per cent of the total income and the top ten per cent received 61.1 per cent of the total income when the Gini coefficient was estimated to 0.67.

In 1989 another effort was made to investigate the distribution of income in the population. *Table 7* shows that the population then was divided into two groups; the subsistence sector and the modern sector, where the latter sector is sub-divided into two subgroups.

[10] Since there mainly seem to be small surveys that have been undertaken, it is possible that some previous studies regarding income distribution in Namibia have not come to my knowledge.

Table 7 Income in subsistence and modern sector 1988

Sector	Share of GDP (per cent)	Share of population (per cent)	Annual per capita GDP (US$)
Subsistence sector	3.4	54.8	85
Modern sector	97.6	45.1	2 531
blacks	25.4	40.0	750
whites	72.2	5.1	16 504

Note: The first two columns do not sum to exactly 100 per cent.

Source: World Bank (1992).

About 45 per cent of the population are employed in the modern sector, and about 55 per cent are employed in the traditional (subsistence) sector. There is a huge difference in per capita income between the two sectors as well as between the population groups. The white population has an annual per capita gross domestic product (GDP) of about US$ 16 500, while the black majority has an annual per capita GDP of US$ 750 in the modern sector and only US$ 85 in the traditional sector. The table further indicates that the Namibian economy is of a dual nature. An overall estimate of GDP per capita is a rather misleading welfare indicator, since it hides the great income differences between the population groups in these two economies.

Another study, see *Table 8*, shows the cash income shares of the bottom 40 per cent and the top 20 per cent of the population in Katatura and Owambo (Owamboland).

Table 8 Share of income distribution in selected areas 1990

Area	Bottom 40%	Top 20%
Katutura	12.8	46.2
Peri-urban Owambo	5.8	65.9
Rural Owambo	4.7	67.4
Total sample	**4.5**	**65.2**

Source: World Bank (1992).

Here Katatura is defined as the black township of the capital Windhoek and the area Owambo consists of the four regions Ohangwena, Omusati, Oshana and Oshikoto, all in the northern part of Namibia. This table suggests that the income inequality is greatest in Owambo, particularly in the rural area. This

result mainly depends on the character of the income sources for the population in rural Owambo, where most of the population lives on subsistence farming and thus not much cash income are generated.

The bottom 40 per cent of the population in Katutura, peri-urban Owambo and rural Owambo in 1990 had an average cash income of 252, 104 and 27 Rands per month, respectively. The corresponding figures for the top 20 per cent of the population are 1 460, 986 and 425 Rands per month, respectively (World Bank, 1992). These income figures, which are in cash, indicate that people in rural Owambo are poorer than the people are in Katatura. This is due to the fact that there are more cash income employment opportunities in geographically well-located Katatura.

A study carried out in January 1992 shows that there are income variations across regions, due to the regional segmentation in the economy (University of Namibia, 1995). For example, the monthly income in Owambo is N$ 683 and in Caprivi it is N$ 572, where crop farming is the main agricultural activity. Part of Omaheke at the border to Botswana and part of Kunene at the border to Angola had a monthly income of N$ 1 361 and N$ 827, respectively. In these two regions cattle rearing are the main agricultural activity. Different sources of income in the agricultural sector, and the region in which one lives seem to be important determinants of the income inequality in Namibia.

4 Measuring and decomposing income inequality

Measuring incomes in developing countries is difficult and therefore the first part of this section considers this issue. The second part considers the choice of adjusted income concept; total household income, per capita income and three versions of adult equivalent income are compared. The third part deals with the measurement of income inequality and the decomposition of total inequality into within-group inequality and between-group inequality.

4.1 Income measurement in developing countries

The collection of income statistics has many conceptual and practical problems in developing countries. One of the first difficulties is that incomes are partly or fully received in kind. Further, commodities produced in the household for

self-consumption do not enter the market, and thus do not involve any money transactions. In both cases imputed prices must be used. Already at this point errors appear, because of the difficulties to impute "true" values. Another problem arises in the estimation of incomes from the agricultural sector, which usually occupies a large proportion of the labour force. The flow of income in this sector is very erratic across seasons, and it is therefore important to use data for the entire year in the construction of income estimates. In this context, measurement errors are likely to arise because of incorrect recollection.

In addition, price disparities exist between urban and rural areas, and hence money income differences may not coincide with real income differences. Defining income could also be difficult. For example, money received as a gift for non-economic reasons is not an income. Since people in developing countries do not record their income transactions to the same extent as people in developed countries do, it is often difficult to separate gifts from regular incomes.

Problems are also associated with the design of questionnaires used in income surveys. The questions should be asked in such a way that the respondent is able to answer without difficulties. However, many individuals have difficulty answering questions such as "What is your income from your own business?".

The unwillingness to answer questions relating to income is also a problem. Some poor people may not want to reveal how poor they are, and may thus overstate their income. The contrary may be the case for more wealthy individuals, who may understate their income for fear of higher tax liabilities.

Many developing countries record consumption expenditures rather than income in household surveys, because of the problems described above. Indeed, consumption may well be a more relevant basis upon which to decide economic status, since income may be temporarily high or low, whereas consumption behaviour may be smoothed across seasons by savings and dissavings (Fields, 1994).[11]

4.2 The choice of adjusted income concept

In general, per capita income is preferred to total household income, because the per capita income measure takes household size into account (Datta &

[11] For further discussion of income measurement see Bigsten (1987).

Meerman, 1980). In addition, adult equivalent income is preferred to per capita income, since the per capita income measure does not account for household composition. Adult equivalence scales recognise that. For example, a four person household cannot live as cheaply as a one person household, but in contrast a four person household does not need four times the resources of a one person household in order to reach the same level of economic welfare.

There is no official adult equivalence scale explicitly made for Namibia, but there is one scale that is used by the Central Statistics Office (CSO) in the context of poverty line (CSO, 1995b). Since this scale was adjusted from a poverty datum line from Botswana I believe it is important to do a sensitivity test of the choice of adult equivalence scales.[12] There is a wide range of different adult equivalence scales that one could test, and the choice is not self-evident.

In *Table 9* income distribution measures based on total household income and per capita income are compared with three kinds of adult equivalence scales that have been taken from different sources. Alm Stenflo (1992) uses the first equivalence scale in the case of poverty in urban Zimbabwe.[13] The second scale is commonly used in publications for OECD countries (Atkinson *et al.*, 1995). Bigsten & Makonnen (1996) also use the second adult equivalence scale. The third scale is the one that is used by the CSO (CSO, 1995b).

[12] See Coulter *et al.* (1992) for a discussion about changing the weights of equivalence scale.
[13] See also Makonnen (1993).

Table 9 Income distribution shares and Gini coefficients by total household income, per capita household income and adult equivalent income, 1993/94

Decile	Total household income	Per capita household income	Adult equivalent income		
			Scale 1	Scale 2	Scale 3
Decile 1	0.8	0.5	0.7	0.6	0.6
Decile 2	1.4	0.9	1.2	1.1	1.0
Decile 3	1.9	1.3	1.6	1.5	1.4
Decile 4	2.5	1.7	2.1	2.0	1.9
Decile 5	3.3	2.3	2.7	2.6	2.5
Decile 6	4.4	3.1	3.6	3.4	3.3
Decile 7	6.0	4.5	5.0	4.8	4.7
Decile 8	9.0	7.0	7.6	7.4	7.2
Decile 9	16.3	14.2	15.1	14.8	14.6
Decile 10	54.4	64.4	60.6	61.8	62.8
Mean income (N$)	16 236	4 753	6 587	6 094	5 184
Gini coefficient	0.6560	0.7334	0.7016	0.7118	0.7197

Notes: See Section 5 and Appendix D for the computation of the income measure. The exchange rate in December 1993 was 3.65 N$/US$. Scale 1: Head of household = 1.0, All other adults = 0.7, Children under 15 years = 0.3; Scale 2: Head of household = 1.0, All other adults = 0.7, Children under 15 years = 0.5; Scale 3: All adults = 1.0, Children between 6 and 15 years = 0.75, Children 5 years and below = 0.5.

The Gini coefficient for total household income is about 0.66 and for per capita income about 0.73, while adult equivalent income falls in between these income measures being 0.70, 0.71 and 0.72, respectively. Thus, it seems to be important to both take household size and composition into account. A closer look at the Gini coefficients and the income shares, indicate that the choice of weighting procedure in our case has a minor influence, since the results are roughly the same regardless of which adult equivalence scale is used. The particular choice of the adult equivalence scale is therefore of little consequence. In fact, there is at most a 2.6 per cent increase of the Gini coefficient going from one scale to another. The second scale is chosen and the reason is that since the three scales are almost identical one can always take the one in the middle.

4.3 The income inequality measurement and the decomposition

In the literature there exist various indices of income inequality (see e.g. Nygård & Sandström, 1981). The presentation is based on the Gini coefficient

and income shares, which both can be derived from the Lorenz curve. The Lorenz curve shows the relationship between the cumulative percentage of households or individuals, ordered from poorest to richest, and the cumulative percentage of total income that accrues to the same households/individuals in the population. To investigate whether total inequality is due to within-group inequality or between-group inequality I use the additively decomposable inequality measures proposed by Theil (1967), the Theil entropy index T and Theil's second measure L.

The Gini coefficient is a measure that can take values between zero (0) and one (1), where 0 implies no inequality and 1 complete inequality. The Gini coefficient (G) can be written as follows:

$$G = \frac{1}{2N^2\mu}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|y_i - y_j\right| \qquad (1)$$

where N in my case is the household population, μ is the mean household income, and y_i and y_j denote the incomes of units i and j. This is the definition of the Gini coefficient that was proposed by Kendall & Stuart (1963).

The disadvantage of the Gini coefficient is that it is not additively decomposable (see Appendix B in Anand, 1983). An additively decomposable inequality index is a measure that can be written as a sum of inequality within-groups and inequality between-groups. In order to determine the within-group inequality and between-group inequality I make use of both the Theil entropy index T and Theil's second measure L. The weights used to calculate these two inequality measures are the income shares of the subgroups in the case of Theil's entropy index T while in the case of Theil's second measure L the population shares of the subgroups is used.

Theil's entropy index T and Theil's second measure L are given by

$$T(y,N) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i}{\mu}\right)\log\left(\frac{y_i}{\mu}\right) \qquad (2)$$

and

$$L(y,N)=\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\mu}{y_i}\right) \qquad (3)$$

where N is the household population, μ the household mean income and y_i is the household income.

When the household population is divided into non-overlapping aggregated categorical subgroups, the Theil's entropy index T and Theil's second measure L can be decomposed into within-group inequality and between-group inequality according to

$$T(y,N)=\sum_{g=1}^{G}T(\mathbf{y}_g,N_g)\left(\frac{\mu_g N_g}{\mu N}\right)+\frac{1}{N}\sum_{g=1}^{G}N_g\left(\frac{\mu_g}{\mu}\right)\log\left(\frac{\mu_g}{\mu}\right) \qquad (4)$$

and

$$L(y,N)=\sum_{g=1}^{G}L(\mathbf{y}_g,N_g)\left(\frac{N_g}{N}\right)+\sum_{g=1}^{G}\left(\frac{N_g}{N}\right)\log\left(\frac{\mu}{\mu_g}\right) \qquad (5)$$

where g is the number of non-overlapping subgroups (g = 1,…, G), $\mathbf{y}_g$ is the vector of incomes in subgroup g, N is the household population, N_g is the household population in the g*th* subgroup, μ is the household mean income and μ_g is the mean income of subgroup g. For both equation 4 and equation 5 the first term on the right hand side represents the within-group inequality and the second term represents the between-group inequality.

5 The data and the income concept employed

Household data from the 1993/1994 Namibia Household Income and Expenditure Survey (NHIES) are used. This survey was done by the Central

Statistics Office (CSO) within the National Planning Commission, Windhoek, during November 1993 to October 1994.[14] The data contains information about consumption and expenditure among Namibian private households.[15] For further details on the design of the data see Appendix C.

The data comprise 4 397 households, 1 712 in urban areas and 2 685 in rural areas. I use most of the information collected through this survey. The data contain information on household composition and size, the main source of income for the household, the primary language spoken, sex and level of educational attainment of head of household, region, urban/rural residence and the household's access to grazing land.[16] Primary language spoken and grazing land will not be included in the analysis. With respect to the former, I do not have a sufficient number of observations for each of the languages spoken in Namibia to perform a meaningful comparison between the different languages. Grazing land gives almost the same results as main source of income, and thus do not contribute much additional information. Total consumption, own produced goods or received in kind, imputed rents, remittances in cash given away, housing and domestic animal investments in cash and in kind, savings and other investments, income tax and other wage deductions, and finally the loan amount are available variables of the household. In addition, remittances in cash received are available, but only in the individual data of the NHIES. Thus, this latter variable had to be aggregated up to household level. It was reported both on a monthly and a yearly basis, but here I make use of the yearly, because remittances usually are made once or twice a year.[17]

It is not possible in this dataset to compute household income by simply adding up income from different income sources. I therefore need to estimate income in a somewhat roundabout way. I build up the estimate of total annual household income by using available information on different types of expenditures and receipts. Total income is set equal to the value of consumption + remittances in cash given away − remittances received + housing investment

[14] In co-operation with Statistics Sweden.

[15] The so-called institutional households, i.e. hospitals, hostels, barracks and prisons are excluded.

[16] See Appendix B for the definitions of the variables used in this study.

[17] The alternative of using the monthly measure multiplied by twelve was considered. This measure became considerable large compared with the yearly and since it is only about 4 per cent of the households in Namibia that depend on remittances as a main source of income this was another reason for choosing the yearly base.

financed by cash + investment in kind + animal investment in cash + savings and other investments including repayment of loans + taxes paid + other non-consumption expenditure – loans taken during the year. (See Appendix D for further details.) The income measure is based on consumption expenditure during a year despite the fact that each household only participated in the survey during one month. For certain items of commodities the annual value was reported. Further, other items that were reported on a monthly base were estimated to annual values by CSO in Windhoek, Namibia. The measure computed here gives a more reliable estimate of income than cash income only. For example, the rural households in Namibia are dependent on own produced commodities during at least certain parts of the year. The estimated proportions of own produced food or food received in kind of total average annual private household consumption is as high as 21.5 per cent in the rural areas.[18] An additional reason for choosing consumption rather than cash income is that the cash income from subsistence farming may fall drastically during drought periods, which are quite common in Namibia. Such fluctuations make cash income estimates more uncertain than consumption estimates.

By using the above-mentioned concept of income, six households came up with negative incomes. Since the computation of the Gini coefficient or Theil's second measure L allows only strictly positive incomes I had either to drop the households or change the negative values to zero. The latter alternative introduces bias in the estimates, and since there were only six households with negative incomes I decided to drop them from the sample. Further, 94 households were excluded due to missing data on the socio-economic variables that I use to investigate the income inequality among the households. Otherwise I include all participating households even the ones with low or high incomes. There were only four such households, two with extremely low incomes and two with extremely high incomes. However, to check if there were any differences in the results I dropped these four households, but the change in the results was minor. By means of the household's location, level of educational attainment, and the main income source it seemed reasonable to include them in the analysis. The final sample constitutes 4 297 households: 1 670 in urban areas and 2 627 in rural areas.

[18] Own calculations based on Table 9.25, p. 198 CSO (1996a).

There are thirteen regions in Namibia, which I aggregate into two major regions: the Central/southern region and the North/north-east region.[19] This aggregation is performed due to the fact that there are too few observations in each region when dividing the sample into the various socio-economic variables and its subgroups. Nevertheless, *Table 10* presents descriptive statistics of the annual household income by Namibia's thirteen regions in order to give a picture of the imbalance of the income between the regions.

[19] The Central/southern region constitutes of Erongo, Hardap, Karas, Khomas, Kunene, Omaheke and Otjozondjupa. The North/north-east region constitutes of Caprivi, Ohangwena, Okavango, Omusati, Oshana and Oshikoto. This is not the same definition that has been used by CSO, Windhoek (CSO, 1996a). The difference is that in this analysis Kunene is included into the Central/southern region, due to more similar characteristics, such as sources of income, geographic segmentation, population size, with the other regions included in the Central/southern region.

Table 10 Descriptive statistics of annual household income (N$) by regions, 1993/94

Region/urban and rural areas	n	Mean income	Median income	Mean/capita income	Min income	Max income
Erongo	371	19 290	10 839	6 140	377	369 454
Hardap	217	20 877	7 121	6 299	599	375 046
Karas	223	25 170	8 865	8 500	480	510 839
Khomas	572	44 271	25 963	13 720	386	1 018 811
Kunene	227	10 357	5 624	2 967	143	150 131
Omaheke	209	16 350	5 839	6 457	300	416 710
Otjozondjupa	359	13 286	7 318	5 187	317	285 022
Central/southern region	*2 178*	*25 018*	*10 027*	*8 121*	*143*	*1 018 811*
Caprivi	228	6 519	3 779	1 636	405	38 620
Ohangwena	415	6 345	3 874	1 028	646	178 361
Okavango	300	8 848	5 940	1 778	258	104 818
Omusati	432	8 201	5 169	1 304	566	98 076
Oshana	401	10 305	5 444	2 063	378	190 117
Oshikoto	343	8 527	4 893	2 241	336	172 085
North/north-east region	*2 119*	*8 141*	*4 741*	*1 649*	*258*	*190 117*
Urban	1 670	30 274	14 092	9 365	300	1 018 811
Rural	2 627	9 092	4 497	2 406	143	510 839
National average	**4 297**	**16 236**	**6 651**	**4 753**	**143**	**1 018 811**

Notes: Sample weights are used when calculating the income statistics except for the first column showing the number of observations in the sample (n). The exchange rate in December 1993 was 3.65 N$/US$.

The variations are due in part to the fact that the varied climate and geographic segmentation in Namibia brings the regions into different income activities, which in turn give differences in the household income. These differences are reflected not only by the geographical segmentation, but also by the differences between the traditional and modern sector. This might particularly be true for the urban/rural difference. One has also to keep in mind that the disparities in prices and cost of living probably differ between the urban and rural areas causing some of the differences. The smallest ratio between mean and median income is observed in the regions of Okavango, Ohangwena and Omusati, which all belong to the North/north-eastern part of Namibia. A contributory factor is that there are no urban residences that bring up the mean, since the two latter regions consist of only rural areas and in Okavango less than one quarter

of the households is situated in urban areas. All regions that belong to Central/southern region, except Kunene and Otjozondjupa, have a mean above the national average. In terms of per capita income Ohangwena is the poorest region of them all, while Khomas, which has the capital, is the outstanding region with the highest per capita income.

6 Empirical results

In this section I present the results of the Gini coefficients and the income shares of the bottom 20 per cent and the top ten per cent of the households. The results from the decomposition of total inequality into within-group inequality and between-group inequality are also presented.

The Gini coefficients and the income shares of the bottom 20 per cent and the top ten per cent of the households, by the two regions Central/southern and North/north-east, are presented in *Table 11*. The head of household is classified besides the regions into six socio-economic variables: urban/rural residence, main source of income, level of highest educational attainment, gender, age and household composition.

Table 11 Gini coefficient, mean income and share of income (%), 1993/94

Subgroups	Central/southern region				North/north-east region			
	Gini coeffient	Mean income (N$)	Income share of Bottom 20 % of pop.	Income share of Top 10 % of pop.	Gini coeffient	Mean income (N$)	Income share of Bottom 20 % of pop.	Income share of Top 10 % of pop.
Region	0.6845	10 245	1.4	54.4	0.5529	2 269	3.8	46.0
Urban/rural residence								
Urban	0.6223	13 347	1.9	46.9	0.6119	5 644	2.2	48.3
Rural	0.7342	6 092	1.7	67.1	0.4852	1 806	4.7	38.9
Main source of income								
Wages	0.6320	10 323	1.9	47.9	0.5879	5 118	2.7	47.5
Business	0.7197	23 538	0.5	52.7	0.5745	2 427	3.7	49.5
Subsistence farming	0.7709	5 355	1.8	73.6	0.4337	1 581	5.3	33.3
Pension	0.7523	6 166	1.5	69.1	0.4832	1 515	4.9	39.5
Cash remittances	0.5792	3 452	3.2	49.4	0.3995	1 357	6.7	33.3
Education of HH								
No formal education	0.5141	2 202	3.8	39.4	0.4192	1 430	5.6	32.2
Primary education	0.5095	3 649	4.0	40.1	0.4822	1 738	4.6	38.3
Secondary education	0.6166	13 978	2.0	47.5	0.5877	3 856	3.1	49.0
Tertiary education	0.4524	33 035	3.3	31.0	0.5033	8 029	3.4	36.7
Gender of HH								
Female	0.6601	7 069	1.8	53.5	0.4596	1 806	5.0	35.6
Male	0.6849	11 470	1.3	53.9	0.6044	2 685	3.1	51.7
Age of HH								
Below 25	0.5184	4 947	3.4	40.0	0.4541	2 656	5.8	39.5
25–34	0.6255	10 555	2.1	48.9	0.5743	3 252	3.3	46.7
35–44	0.6381	10 744	1.6	46.9	0.5744	2 733	3.5	46.9
45–54	0.7334	12 763	1.1	59.1	0.5513	2 034	4.0	46.0
55–64	0.7378	7 334	1.4	63.6	0.5584	1 916	3.8	47.3
65+	0.7737	10 669	0.9	66.2	0.4141	1 419	5.7	31.9
Household composition								
Single	0.6078	13 403	1.9	43.8	0.6132	7 478	2.6	46.3
Single with children	0.5998	7 634	2.0	45.8	0.4472	1 919	4.9	32.9
Single extended family	0.5923	4 571	2.7	47.2	0.4917	1 866	4.6	40.1
Couple	0.7166	26 071	1.0	56.9	0.5260	3 103	5.0	44.0
Couple with children	0.6196	14 655	1.2	42.3	0.6093	2 673	3.4	52.4
Couple extended family	0.6464	5 666	2.0	52.9	0.5127	1 898	4.1	40.5

Notes: Results for the subgroup other income is not reported to too few observations for computing the Gini coefficient. The adult equivalent scale where Head of household (HH) = 1.0; All other adults = 0.7; Children under 15 years = 0.5, is used. The exchange rate in December 1993 was 3.65 N$/US$.

The table suggests that income inequality, measured by the Gini coefficient, seems to be larger in the Central/southern region for all subgroups. Only in the case of a single household is the Gini coefficient slightly larger in the North/north-east region. The Central/southern region and the North/north-east region have a Gini coefficient of 0.68 and 0.55, respectively. The high Gini coefficient in the former region is partly explained by the fact that the region is of a more heterogeneous character, i.e. the economic sectors have a broader mix of activities involving unskilled, semi-skilled and skilled workers. The more heterogeneous the structure of an economic sector, the greater is the income disparities (Lecaillon *et al*., 1984). A reason for the lower Gini coefficient in the North/north-east region is that the majority of population is composed of people living in rural areas. The top 10 per cent of the households have about 54 per cent of the income share in the Central/southern region, while the corresponding figure for the North/north-east region is 46 per cent.

The breakdown by urban and rural areas reflects the differentiation between the non-agricultural (modern) sector and the agricultural (traditional) sector. The Gini coefficient for the rural areas in the Central/southern region is higher than that of the urban areas, which contradicts general findings: income distribution is usually more unequal in urban areas than in rural areas.[20] This is, however, not the case in the North/north-east region. The urban Gini coefficient for both regions is about 0.60, which reflects that even the income disparities can be substantial between the traditional urban sector and the modern urban sector. The droughts that frequently afflict Namibia also contribute to the high urban inequality, because the drought force people in rural areas dependent on agriculture to move to urban areas to look for wage work. This has the effect of increasing urban unemployment, thus widening the urban income gap. Namibia's overall urban Gini coefficient is about 0.63, which is high compared with a study of urban Ethiopia with data from 1994 showing a Gini coefficient of about 0.56 (Bigsten & Makonnen, 1996). The overall Gini coefficient of Namibia is about 0.71, which is a very high Gini coefficient compared with other sub-Saharan African countries. For example, Kenya, South Africa and Tanzania have Gini coefficients of 0.575, 0.584 and 0.381, respectively (World Bank, 1997).

[20] For some empirical evidence see Sundrum (1990), p. 96 where Gini coefficients are estimated for urban and rural areas from nine countries, where only two of them show a slightly higher value of the rural Gini coefficient.

There is a considerable difference in the Gini coefficient between the rural areas of the Central/southern region, which have a Gini coefficient of 0.73 and the rural areas of the North/north-east region, which have a Gini coefficient of 0.49. This difference is essentially due to the co-existence of commercial and subsistence farmers in the Central/southern regions, where the commercial farmers in general have a much higher income than that of subsistence farmers. The two results may also reflect that the traditional sector has grown much faster in the Central/southern region than in the North/north-east region. Rural data from 1991 for Tanzania show a Gini coefficient of about 0.72 (Ferreira, 1996), which is more comparable to my rural Gini coefficient for the Central/southern region than for Namibia's overall rural Gini coefficient of 0.66.

In the Central/southern region, the highest Gini coefficient, according to main source of income, is found in the subgroup subsistence farming with a value of 0.77. The first reason that may explain this is the distinction between commercial farming and communal farming in size, ownership and the use of different techniques. The second reason is the drought during part of the survey year that affected production adversely, which in turn may generate low income for some of the farmers and thus higher inequality in this subgroup compared with the other income groups. In contrast, the North/north-east region has a Gini coefficient of 0.43 for the same subgroup, suggesting a more homogeneous income structure in the agricultural sector. In this region, the subgroups wages and business have the highest Gini coefficient, 0.59 and 0.57, respectively.[21] One possible explanation of the high inequality in the subgroup wages could be the employment opportunities for high paid work in the public and administration services, which was lately established in the region. In the North/north-east region, compared with the Central/southern region, the lower inequality in the subgroup pension may be explained by the fact that households in the latter region to a larger extent have pension from the state, i.e. private pensions. A study of South African income inequality with data from 1993 has estimated Gini coefficients for different income sources (Leibbrandt *et al.* 1996). Their results show a Gini coefficient of 0.655 for wage income, which is rather similar to my result. Furthermore, their result of the Gini

[21] These values are, however, lower compared with the Central/southern region.

coefficient for agriculture is 0.931 and for remittances it is 0.840, which is much higher than in my case.[22]

In the literature, education seems to be the most important determinant of income inequality (e.g. see Nafziger, 1988 or Sundrum, 1990), and is thus of particular interest. In the Central/southern region there does not exist a pattern for the Gini coefficient. It is first high, then low; thereafter it increases and at the tertiary level it decreases. In the North/north-east region the Gini coefficient increases with increased level of educational attainment for the head of the household up to secondary education, whereas it decreases at the tertiary level.[23] Note that in both regions the head of households with secondary education have the greatest inequality. This result may reflect that some of the well-educated heads in this subgroup have low-paid jobs, because of the difficulties of finding a job that corresponds to their level of education. In other words, some persons have had to take a job that did not require their level of education and thus might have had to accept lower payment.[24]

Rather similar results of the Gini coefficients are obtained for the female- and male-headed households in the Central/southern region, 0.66 and 0.68 respectively.[25] The differences in income inequality between female- and male-headed household in the North/north-east region may first of all reflect differences in employment opportunities. For example, women from the northern part of Namibia were historically excluded from the Labour Contract System, and thus from wage work. The results also suggest that women work less outside the household than men do. The Gini coefficient among females is substantially higher in the Central/southern region than in the North/north-east region. This probably reflects that some of the females in the former region to a higher extent work outside the household, not only with agriculture, but also, for example, as teachers, nurses and secretaries. This raises the inequality in this subgroup. The low average income that female-headed households have in the North/north-east region may be explained by the fact that they depend largely on agriculture. They have difficulties in cultivating large areas of land

[22] Some caution in the results for the subgroup remittances in the North/north-east region should be taken, because less than 60 observations are recorded.
[23] The result in the subgroup tertiary education in this region may be inconsistent, because there are less than 60 observations in the sample. The result should, thus, be interpreted with caution.
[24] Note that mean income rises with education.
[25] For the bottom 20 per cent and the top 10 percent of the population the income shares are almost identical.

because they lack help from skilled labour. In turn this has the effect of lowering output, which gives a decrease in average income.

The Gini coefficients in the North/north-east region show Kuznet's pattern of the 'inverted U' between the age of the head of household and income distribution, i.e. low inequality in the young and the old age groups with a peak in the age group 35–44. The great inequality in the older age group in the Central/southern region may in part be a reflection of a more unequal distribution of education between the older generation compared with the younger age groups. Further, in this region the high value of the Gini coefficient in the age group 65 and over seems, at first sight, odd because it is assumed that this group should be more equalised than other age groups, with the exception of the age group below 25. One possible explanation to the result could be that some of the household population in age group 65 and over is still working, while others do not.[26]

In the Central/southern region the highest Gini coefficient is found among couples without children, while in the North/north-east region the highest Gini coefficient is found among single households. For both regions, single households with children and single extended families have the lowest Gini coefficients. The income share of the bottom 20 per cent of the household population is smaller for all subgroups in the Central/southern region compared with the North/north-east region.

Table E1 of Appendix E presents results of estimating the Gini coefficient, when the variable total consumption (TCONS) is used.[27] Comparing the results, the Gini coefficients are lower in *Table E1*. Nevertheless, some subgroups still have a fairly high value of the Gini coefficient, indicating that Namibia is a country with a highly unequal distribution. The most striking differences between the Gini coefficients are found for subsistence farming in the Central/southern region and for business in the North/north-east region.

Table 12 shows the decomposition of total inequality into within-group inequality and between-group inequality using both Theil's (1967) entropy index T and Theil's second measure L. Income and population percentages are also reported in the table.

[26] The highest average incomes are found in the households with a head of an age 45–54 in the Central/southern region, whereas in the North/north-east region the young generation 25–34 has the highest average incomes.

[27] Using the total consumption measure the Gini coefficients for overall Namibia, urban areas and rural areas are 0.67, 0.61 and 0.60, respectively (not shown in *Table E1*).

Table 12 The decomposition of total inequality for income by subgroups, 1993/94

Subgroups	Central/southern region				North/north-east region			
	T	Inc. (%)	L	Pop. (%)	T	Inc. (%)	L	Pop. (%)
Urban/rural residence								
Urban	0.7617	76.5	0.7780	57.2	0.7138	24.4	0.7208	12.1
Rural	1.3103	23.5	1.0745	42.8	0.4851	75.6	0.4093	87.9
Within-group inequality	0.9012		0.9048		0.5537		0.4469	
Between-group inequality	0.0651		0.0709		0.1137		0.0907	
Main source of income								
Wages	0.7646	72.8	0.7995	70.5	0.6654	36.2	0.6602	18.9
Business	1.0825	16.5	1.3650	7.7	0.6651	4.8	0.5780	3.9
Subsistence farming	1.7805	3.1	1.2037	5.1	0.3508	49.5	0.3242	62.1
Pension	1.3445	5.9	1.1713	11.1	0.5714	7.5	0.4066	11.7
Cash remittances	0.6417	1.6	0.6062	5.4	0.3227	2.0	0.2714	3.3
Other income	0.3531	0	0.4752	0.2	0.0253	0	0.0254	0.1
Within-group inequality	0.8844		0.8938		0.5150		0.4053	
Between-group inequality	0.0819		0.0818		0.1525		0.1322	
Education of HH								
No formal education	0.5204	5.9	0.4726	24.6	0.3248	25.8	0.3026	35.0
Primary education	0.4955	9.9	0.4543	24.2	0.4915	31.9	0.4069	40.2
Secondary education	0.7817	57.6	0.7514	42.5	0.6997	32.9	0.6198	22.0
Tertiary education	0.3690	26.6	0.4180	8.7	0.4688	9.4	0.5107	2.8
Within-group inequality	0.6271		0.5819		0.5301		0.4199	
Between-group inequality	0.3392		0.3938		0.1373		0.1177	
Sex of HH								
Female	0.8489	17.7	0.8643	27.8	0.4051	38.2	0.3639	47.4
Male	0.9695	82.3	0.9884	72.2	0.7958	61.8	0.6572	52.6
Within-group inequality	0.9463		0.9539		0.6485		0.5182	
Between-group inequality	0.0200		0.0218		0.0190		0.0194	
Age of HH								
Below 25	0.4987	2.9	0.4925	7.2	0.4241	3.5	0.3483	4.6
25–34	0.7520	25.4	0.7639	26.4	0.6535	22.2	0.6009	18.9
35–44	0.7510	31.4	0.8450	27.1	0.7039	25.0	0.5850	20.1
45–54	1.2313	23.1	1.1936	17.3	0.7258	16.9	0.5350	17.5
55–64	1.1632	8.2	1.1498	10.9	0.7394	14.1	0.5523	14.9
65+	1.3224	9.0	1.3715	11.1	0.3160	18.3	0.2937	23.9
Within-group inequality	0.9441		0.9499		0.6250		0.4937	
Between-group inequality	0.0222		0.0257		0.0425		0.0438	

Table 12 continues The decomposition of total inequality for income by subgroups, 1993/94

Subgroups	Central/southern region				North/north-east region			
	T	Inc. (%)	L	Pop. (%)	T	Inc. (%)	L	Pop. (%)
Household composition								
Single	0.6520	7.3	0.7615	13.5	0.7005	4.0	0.7084	4.3
Single with children	0.6800	3.9	0.7061	5.9	0.3438	6.0	0.3459	9.8
Single extended family	0.6881	15.5	0.6456	26.5	0.5021	34.7	0.4198	38.3
Couple	1.0852	14.9	1.1701	8.3	0.5579	1.5	0.4703	2.4
Couple with children	0.6907	37.2	0.8793	22.3	0.8487	15.9	0.6580	14.5
Couple extended family	0.8811	21.2	0.8121	23.4	0.5163	37.9	0.4635	30.8
Within-group inequality	0.7914		0.7997		0.5816		0.4741	
Between-group inequality	0.1749		0.1760		0.0859		0.0635	
Total inequality	**0.9663**		**0.9757**		**0.6675**		**0.5376**	

Note: The adult equivalent scale, where Head of household (HH) = 1.0, All other adults = 0.7, Children under 15 years = 0.5, is used.

The results suggest that it is the within-group inequality that is the major component in total inequality. The between-educational contribution to total inequality in the Central/southern region is as high as 35.1 per cent and 40.4 per cent using T and L, respectively. Education gives the highest between-group contribution to total inequality of all socio-economic variables in this region. Equalising the between-group inequality would reduce the total inequality from about 0.97 to about 0.63. In the North/north-east region the corresponding between-educational contribution to total inequality is only 20.6 per cent using T and 21.9 per cent using L, but still the second highest contribution of the between-group inequality to total inequality compared with the between-group contribution of the other socio-economic variables. These results suggest that the between-educational contribution to total inequality is of great importance, particularly in the Central/southern region. A comparative study on Lesotho using data from 1986/87 also finds that the between-educational contribution to total inequality is the most important one of all variables analysed (Makonnen, 1993).

In the Central/southern region the lowest between-group contribution, about 2 per cent using either T or L, is received in the socio-economic variable gender of head of household. This is the case in the North/north-east region as well, but here the figures are somewhat higher, i.e. 2.8 per cent and 3.6 per cent using T and L, respectively. In both regions the male-headed households have

more inequality compared with the female-headed households. In the North/north-east region the male-headed have inequality rates twice as high as their counterparts. This suggests that people in female-headed households work more inside the traditional sector than do people in male-headed ones.

Inequality increases with age of the head of household in both regions, with three exceptions. In Central/southern region the between-age inequality is of no importance, since the results suggest that total inequality would only decrease, using either of the measures T or L, with about 2.5 per cent if equalising between-age inequality. The result of the between-age inequality in the North/north-east region is more than twice as important, since equalising between-age inequality the total inequality would decrease with 6.4 per cent and 8.2 per cent using T and L, respectively.

The second highest contribution of the between-group inequality to total inequality in Central/southern region is found for the socio-economic variable household composition. The contribution is about 18 per cent of total inequality for both T and L. Thus, equalising the between-group inequality would reduce total inequality from about 0.97 to about 0.79. Highest inequality is found among couples, suggesting that for some of the couples both are working, while within other couples there is only one person working. For the same socio-economic variable in the North/north-east region, the between-group inequality is contributing less to total inequality, about 13 per cent using T and about 12 per cent using L. In this region couples with children have the highest inequality. The within- and between-group inequality contribution to total inequality does not change much if the measure of total consumption is used; see *Table E2* in Appendix E. Worth noticing is, however, that total inequality in the North/north-east region is higher using total consumption than the income concept.

7 Summary and conclusion

In this paper the distribution of income among the Namibian households has been analysed. The analysis has been conducted using the 1993/1994 Namibia Household Income and Expenditure Survey. The household's income concept is adjusted by adult equivalence scales. To measure the distribution of income the Gini coefficient, the income shares of the bottom 20 per cent and the top 10 per cent of the household population were calculated. I also investigate whether

total income inequality originates from within-group inequality or from between-group inequality. Namibia's thirteen regions are aggregated into two major regions, one containing the regions in the central and the southern Namibia and the other region containing the regions in the north. Each of the two regions was divided into six socio-economic variables where each variable was further divided into various subgroups.

The results indicate a highly unequal distribution of income in Namibia, particularly in Central/southern region. In this region the Gini coefficient for the various subgroups varied between 0.51 and 0.78, while the corresponding figures for the North/north-east region are much smaller, varying between 0.40 and 0.64. The average income is also much lower for all household categories in the latter region than in the former region. The bottom 20 per cent of the household population in Central/southern region acquire 1.4 per cent of total income in the region, while the top 10 per cent of the household population acquire 54.4 per cent of total income. The corresponding figures for the North/north-east region are 3.8 per cent and 46 per cent.

Examining the decomposition of inequality into within-group inequality and between-group inequality shows that the within-group inequality is the major component in total inequality. Rather different results are obtained from the two regions regarding the contributory component of within- and between-group inequality to total inequality. The highest between-group component to total inequality is found for education in the Central/southern region whereas in the North/north-east region main source of income gives the highest between-group contribution to total inequality. However, the between-education inequality is also high in the North/north-east region. The general presumption, that most of the income inequality in developing countries originates from disparities between urban and rural areas, does not hold in the Namibian case. Equalising the between-urban/rural inequality will only reduce total inequality by 6.5 percentage points in the Central/southern region. The corresponding figure is higher (11.4 percentage points) in the North/north-east region, but still fairly low.

Two main conclusions can be made from this analysis of income distribution in Namibia. First, the Central/southern region suffers from a more skewed distribution of income than the North/north-east region. Second, it is the within-group inequality that is causing the unequal income in Namibia.

In regard to policy intervention, the on-going investments in the area of education must continue. However, further education and training for people

already at work, particularly in small-scale firms and in the informal sector, are also important to consider since most of the population is or will be involved in these areas of the labour market. Improvements of the access to the credit market for small-scale firms as well as for the informal sector would reduce barriers to expand the business. The small-scale firms and people in the informal sector have little access to credit, and thus access to more credit may improve their income and thus reduce the overall income inequality. Further, the results suggest that Namibia has to focus on policies concerning the income disparities between the thirteen regions. However, it will be difficult to apply a general policy for each region, because of the differences in the socio-economic development in the regions.

References

Alm Stenflo, G (1992) “Poverty in Zimbabwe 1990–91, A preliminary report”, Statistics Sweden, International Consulting Office, Stockholm.

Anand, S (1983) *Inequality and Poverty in Malaysia, Measurement and Decomposition*, Oxford University Press, A World Bank Research Publication.

Atkinson, A B, L Rainwater & T M Smeeding (1995) “Income Distribution in OECD Countries, Evidence from the Luxembourg Income Study”, *Social Policy Studies*, No. 18, OECD.

Bigsten, A (1987) “Poverty, inequality and development” in Gemmell, N (ed.), *Surveys in Development Economics*, Basil Blackwell, London.

Bigsten, A & N Makonnen (1996) “The Anatomy of Income Distribution in Urban Ethiopia”, Development Economics Working Paper 1996/1, Department of Economics, School of Economics and Commercial Law at Göteborg University, Göteborg.

CSO (1995a) *1991 Population and Housing Census, Basic analysis with highlights*, SCB-Tryck.

CSO (1995b) *The distribution of economic resources in the population of Namibia, Some highlights*, National Planning Commission, Windhoek.

CSO (1995c) *Women and Men in Namibia*, National Planning Commission, Windhoek.

CSO (1996a) *Living Conditions in Namibia, Basic description with highlights*, National Planning Commission, Windhoek.

CSO (1996b) *National Accounts 1980-1995*, National Planning Commission, Windhoek.

CSO (1996c) *The 1993/1994 Namibia Household Income and Expenditure Survey NHIES, Administrative and technical report*, National Planning Commission, Windhoek.

Coulter, F A E, F A Cowell & S P Jenkins (1992) "Equivalence Scale Relativities and the Extent of Inequality and Poverty", *Economic Journal*, vol. 102, no. 414, pp. 1067–1082.

Datta, G & J Meerman (1980) "Household income or household income per capita in welfare comparisons", *Review of Income and Wealth*, vol. 26, no. 4, pp. 401–418.

Ferreira, M L (1996) "Poverty and Inequality During Structural Adjustment in Rural Tanzania", The World Bank, Policy Research Working Paper no. 1641.

Fields, G S (1994) "Poverty and Income Distribution Data for measuring poverty and inequality changes in the developing countries", *Journal of Development Economics*, 44, pp. 87–102.

Kendall, M G & S A Stuart (1963) "The Advanced Theory of Statistics", vol. 1, *Distribution Theory*, Second Edition, Griffin, London.

Lecaillon, J *et al.* (1984) *Income distribution and economic development, An analytical survey*, International Labour Office, Geneva.

Leibbrandt, M V, C D Woolard & I D Woolard (1996) *The Contribution of Income Components to Income Inequality in South Africa, A Decomposable Gini Analysis*, Working Paper no. 125, World Bank, Washington D.C., Living Standards Measurement Study.

Makonnen, N (1993) *Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho*, dissertation thesis, Economics Studies, no. 40, Department of Economics, School of Economics and Commercial Law, Göteborg University.

Murray, R (1993/94) “Namibia”, *The World of Information, Africa Review*, Kogan Page & Walden Publishing, Seventeenth edition, pp. 126–130.

Nafziger, E W (1988) *Inequality in Africa: Political Elites, Proletariat, Peasants and The Poor*, Cambridge University Press, Cambridge, African Society Today.

National Development Plan 1 (1995) National Planning Commission, Windhoek, draft.

Nygård, F & A Sandström (1981) *Measuring income inequality*, Almqvist & Wiksell International, Stockholm.

Paukert, F & D Robinson (1992) *Incomes policies in the wider context: wage, price and fiscal initiatives in developing countries*, International Labour Office, Geneva.

Sparks, D L (1996) “Economy”, in *Africa South of the Sahara 1997*, Europa Publications Limited, London, Twenty-Sixth Edition, pp. 687–691.

Sundrum, R M (1990) *Income distribution in less developed countries*, Routledge, London and New York.

Theil, H (1967) *Economics and Information Theory*, North Holland, Amsterdam, Studies in Mathematical and Managerial Economics.

UNDP (1996) *Human Development Report 1996*, United Nation, New York.

United Nation (1986), *Namibia: Perspectives for National Reconstruction and Development*, United Nations Institute for Namibia, Lusaka.

United Nation Industrial Development Organization (1990) “Namibia Industrial development at independence”, Industrial Development Review Series, Vienna.

University of Namibia (1995) *Namibia Poverty Profile, A Report for SIDA*, Social Sciences Division Multi-Disciplinary Research Centre, Windhoek.

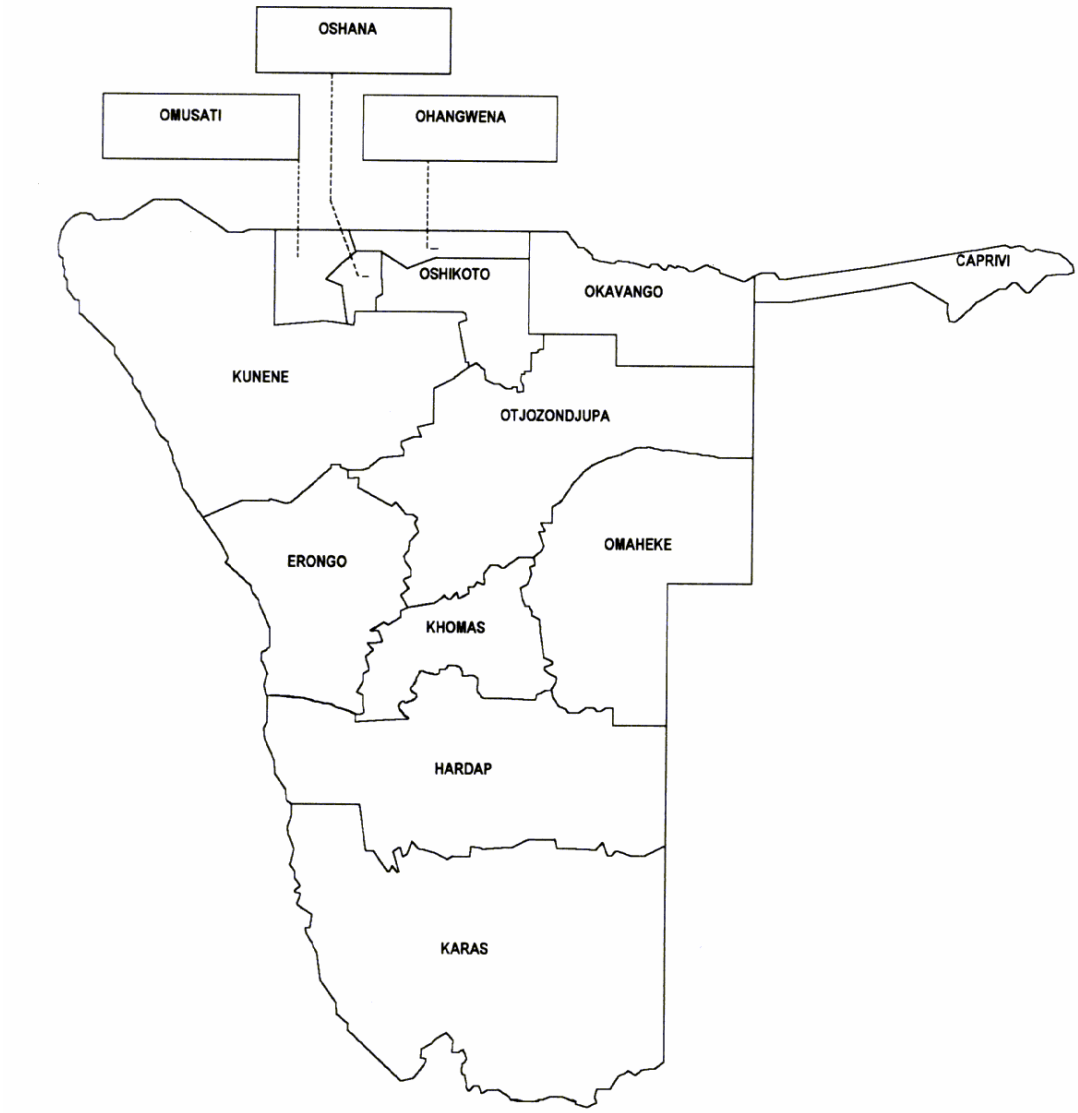
van Ginneken, W (1986) “Incomes and Wages in Namibia”, International Labour Office, Lusaka, Zambia, World Employment Programme, Southern African Team for Employment Promotions (SATEP), Working Paper.

World Bank (1992) *Namibia Poverty Alleviation with Sustainable Growth*, A World Bank Country Study, The World Bank, Washington D.C.

World Bank (1996) *World Development Report 1996*, Oxford University Press, New York.

World Bank (1997) *World Development Report 1997*, Oxford University Press, New York.

Appendix A – A map over Namibia



Source: CSO (1995a).

Appendix B – Definitions of variables

Economically inactive population: Persons aged 15 years and over that are not in the category economically active.

Student: A person who was engaged in studies for the last seven days, in any regular educational institution public or private for systematic instruction at any level of education.

Homemaker: A person who were strictly engaged in household duties during the last seven days, in his/her own home like housekeeping, raising his/her own children, taking care of the old.

Income recipient: A person receiving some money income such as rents, interest from made investments, without actually engaging himself/herself in the related activities.

Disabled, old, retired and others: The two first categories are persons who receive public aid or private support and are living mainly on these receipts. Retired persons receive some sort of pension and are living mainly on these receipts. Others are persons not falling into any of the mentioned categories.

Economically active population: The labour force, i.e. persons employed, underemployed or unemployed.

Employed: A person working for pay, profit or family gain for at least one hour during the seven-day period before the interview or who did not work during that period, but had a job/business to go back to.

Fully employed: Employed person who was not available for more work.

Underemployed: Employed person who was available for more hours of work during the seven-day period.

Unemployed: A person who did not work during the seven-day period before the interview or had no job/business to go back to during that period, but who was available and looked for work during that period.

Paid employee: A person employed for wages, salaries, commission, tips or a payment in kind. Working either on a regular or casual basis, temporary or permanent basis, either on time rate or on piece rate by an employer who, for example, may be a government department or a private firm.

Employer: A person operating his/her own business or a business owned partially along with other partners with the aid of one or more paid employees.

Own account worker: A person operating his/her own business or a business owned partially along with other partners or who works for family gain without aid of paid employees, but can have unpaid family workers.

Unpaid family worker: A person who, for example, works in a household business operated by a related member of the person's household without any payment.

A private household: A private household may consist of one or several persons. A one-person household is a person living alone and catering for himself/herself. A multi-person household contains two or more individuals (relatives or non-relatives), who live together and have common catering arrangements. The definition of a household member is a person who slept in the household during the last 24 hours before an interview and who had common catering arrangements with the household during these 24 hours.

Urban/rural residence: The urban residences are the following 15 municipalities and 12 towns in Namibia.
Municipalities: Swakopmund, Windhoek, Gobabis, Grootfontein, Karibib, Karasburg, Keetmanshoop, Mariental, Okahandja, Omaruru, Otavi, Otjiwarongo, Outjo, Tsumeb, Usakos and Walvis Bay (except the area of the Topnaars).
Towns: Hentiesbaai, Lüderitz, Okakarara, Ondangwa, Ongwediva, Opuwo, Oshakati, Rehoboth, Rundu, Katima Mulilo, Khorixas and Arandis.
The other parts of Namibia are defined as rural residences including the Topnaars in Walvis Bay.

Main source of income: The question "What income source is the most important for the wellbeing of the entire household?" has been the base of how to classify the households into this categorical variable.

Education: The level of highest educational attainment.
No formal education: No school experience or currently in Grad 1.
Primary education: Grade 1–6.
Secondary education: Grade 7–12.
Tertiary education: Diploma or certificate issued by a university, Technikon or Technical institutes; university courses leading to first degrees 1–4 years; post-graduate (diploma, master degree, doctorate degree); teacher training.

Household composition:
Single: A one-person household.
Single with children: A one-person household with one or more children.
Single with extended family: A one-person household, which includes own children's spouse, children's children, parents (including spouse's parents) or other relatives. Own children may or may not be include. Non-relatives are also included.
Couple: A two-person household.
Couple with children: A two-person household with one or more children.
Couple with extended family: A two-person household, which includes own children's spouse, children's children, parents (including spouse's parents) or other relatives. Own children may or may not be include. Non-relatives are also included.

Appendix C – Design of the data

Each household participated in the survey during one full month and recorded daily all their transactions into a record book. The original sample covered 4 752 households, but due to non-response or answers too poor to analyse, the final data from CSO covers 4 397 households of which 1 712 were in urban areas and 2 685 in rural areas. These households comprise 24 984 individuals, which form the individual data. In addition to characteristics such as education, employment etc., individuals also reported cash incomes, i.e. gross wages and allowances; income taxes and other wage deductions; pensions, and household business income, interest, royalties and dividends, as well as cash gift/remittances received.

The sampling of the households has been done by a two-stage sampling procedure for Namibia excluding Walvis Bay. In the beginning of the survey Walvis Bay was still part of South Africa. The Primary Sampling Units (PSU's) were created from a frame of geographical areas. These geographical areas were created from the Census' enumeration areas (EA). To get sufficiently large PSU's, small EA's were combined with neighbouring EA's. The geographical areas contained 80 to 200 households. 192 PSU's, 81 in the urban areas and 111 in the rural areas were sampled.[28] The second stage of the survey sampled the households, the Secondary Sample Unit. They were listed, and a systematic equal probability random sample procedure was used to sample 24 households in each PSU's. The seasonality of the households' consumption and expenditure was taken care of in the sampling, since the sample was divided into 12 monthly equal-sized sub-samples.

On March 1, 1994 Walvis Bay was integrated into Namibia, but for logistic planning reasons Walvis Bay was only included in the NHIES during the last six months of the survey year, i.e. from May 1994. The sample of Walvis Bay is part of the Erongo region. The households for Walvis Bay were selected primarily by a stratified one-stage sampling procedure, since municipality registers of the households already existed. For five households a two-stage sampl-

[28] The sample was basically proportionally allocated, but has a slight over-sampling in the urban areas and in the Omaheke region. The reasons for this are that the income level variation seemed to be larger between the households in the urban areas than in the rural areas, lower survey costs in the urban areas, and that each region should contain at least 10 PSU's in order to receive reasonably good statistics from each region.

ing procedure had to be used, since more than one household stayed in the sampled household, so-called backyard squattering. There also existed areas or population groups that were not covered by the registers and for these household lists were constructed. Unfortunately, some hostel areas had to be excluded for security reasons. For more details and information of the sampling plan see CSO (1996c).

The NHIES sample is provided together with probability weights, which vary across PSU's. Application of these weights makes the sample representtative at the national level. The empirical analyses in this paper employ these weights.

Appendix D – The income concept

Using expenditure variables including the transactions in kind, annual income for each household can be calculated. However, two problems arose when these computations were made. First, there were very high values for some households for investments during the last twelve months. This impelled me to do a careful analysis of the investment variable, to try to find out to what extent the household's investments had been made with borrowed money. Thus, an investigation of the loans for each household was performed, which showed that some households had loans that exactly or nearly corresponded to the invested amount. I decided to adjust for this. I then faced a second problem, namely to find out what the amount borrowed was. The loan is just reported as a stock, and may therefore refer to previous years and not only to the year of the survey. This forced me to make some assumptions: i) if the household had no investments during the year, I assumed that a positive loan amount indicates that the loan had been taken during previous years and the amount borrowed during the year takes the value zero; ii) if the loan amount is greater than the investments I assumed that the loan taken this year equals the investments; and iii) if there is no information about any loan then no adjustment is made to the investments. The reasons for choosing this approach is that I at least, to some extent, alleviate the overestimation of the household income, but might still overestimate the income for the households that have not answered on the question regarding loans. The total annual household income is established as follows:

TOTAL ANNUAL HOUSEHOLD INCOME = TCONS + CREMITG – CREMITR + CINV + KINV + CDMANIM + CSAVE + CTAX + CNONCEXP – LOANAMNT

Total consumption (TCONS): Consists of food expenditure, which include bread and cereals; meat; fish; milk, cheese and egg; oil and fat; vegetables and potatoes; fruits and nuts; sugar; non alcoholic beverages; other food; meals. Own produced food or food received in kind includes bread and cereals; meat; fish; vegetables and potatoes; fruits and nuts. Other consumption items include total clothing and footwear; total housing, i.e. imputed rent in cash and kind, fuel, power, etc.; total furniture and utensils; total household operations; total medical care; total transport and communication; total education; total personal

care; total recreation, i.e. cultural, entertainment; total other, i.e. insurance legal aid.

Remittances in cash given away (CREMITG): This includes ad hoc and regular gifts, remittances and maintenance payments to relatives and friends. Not included are contributions to cultural societies and common village equipment.

Remittances in cash received (CREMITR): The aggregation of individual remittances received within and outside Namibia.

Housing investments financed by cash (CINV): This item consists of cash investments in building materials, land and housing.

Investments in kind (KINV): This item consists of kind investments in building materials, cattle, donkeys, horses, goat, sheep, pigs, poultry and animal care received by the household.

Domestic animal investments in cash (CDMANIM): Cattle; donkeys, horses; goat, sheep, pigs, poultry and animal care.

Savings and other investments (CSAVE): Fees for life/pension insurance; loan and credits given away; repayment of loans and credits (car and mortgages along with clothing and furniture repayments are included as well as interest paid); saving deposit.

Income tax and other wage deductions (CTAX): This item also includes pension scheme and medical aid deductions.

Non-consumption expenditure (CNONCEXP): This item consists of bride prices or lobbola given away and fines or penalties.

Loan amount (LOANAMNT): The household's total borrowed amount during the year.

For the variables of housing, land and animals, an imputed value has been implemented in the following way. If the household has answered that they bought, for example, land during the last twelve months, but did not state the amount then a value is imputed based on corresponding values paid by households in the same or neighbouring PSU. For a more detailed description of the variables see CSO, (1996c).

Appendix E – Consumption distribution

Table E1 Gini coefficient, mean consumption and share of consumption (%), 1993/94

Subgroups	Central/southern region				North/north-east region			
	Gini coeffient	Mean consum. (N$)	Consum. share of Bottom 20 % of pop.	Consum. share of Top 10 % of pop.	Gini coeffient	Mean consum. (N$)	Consum. share of Bottom 20 % of pop.	Consum. share of Top 10 % of pop.
Region	0.6509	7 754	1.8	51.6	0.5174	2 021	4.4	42.8
Urban/rural residence								
Urban	0.5953	10 148	2.3	44.9	0.5835	4 453	2.8	3.2
Rural	0.6801	4 549	2.3	60.7	0.4624	1 687	5.1	37.2
Main source of income								
Wages	0.5968	7 732	2.3	45.4	0.5718	4 171	3.2	46.8
Business	0.7123	18 355	0.6	53.5	0.4977	1 915	4.7	41.3
Subsistence farming	0.6767	3 527	2.7	62.5	0.4179	1 513	5.7	32.2
Pension	0.6998	4 588	1.9	45.4	0.4658	1 454	5.3	38.1
Cash remittances	0.5753	3 534	3.3	62.5	0.3887	1 393	7.2	32.6
Education of HH								
No formal education	0.4870	1 988	4.2	37.0	0.4007	1 361	6.1	31.1
Primary education	0.4790	3 136	4.5	37.5	0.4550	1 633	5.1	36.0
Secondary education	0.5794	10 259	2.6	44.2	0.5619	3 268	3.6	46.7
Tertiary education	0.4568	24 621	3.5	30.6	0.5083	6 139	3.9	36.5
Gender of HH								
Female	0.6335	5 797	2.2	51.1	0.4364	1 698	5.5	34.3
Male	0.6510	8 509	1.8	51.0	0.5675	2 312	3.6	48.0
Age of HH								
Below 25	0.4795	4 272	3.9	37.0	0.4146	2 473	6.4	35.5
25–34	0.5940	8 027	2.6	47.1	0.5325	2 821	4.2	43.4
35–44	0.6008	7 932	2.1	44.5	0.5282	2 282	4.2	42.4
45–54	0.7080	9 600	1.3	56.4	0.5256	1 854	4.5	44.3
55–64	0.7066	5 746	1.7	60.1	0.5467	1 803	4.0	46.4
65+	0.7357	8 027	1.2	61.1	0.3931	1 340	6.1	30.2
Household composition								
Single	0.6004	11 399	2.1	44.5	0.5891	6 549	3.1	45.8
Single with children	0.5620	6 322	2.5	42.8	0.4318	1 875	5.4	31.9
Single extended family	0.5485	3 667	3.3	42.7	0.4625	1 720	5.2	37.8
Couple	0.6744	17 785	1.5	52.2	0.5089	2 981	5.3	43.0
Couple with children	0.5946	10 624	1.6	40.8	0.5467	2 184	4.2	45.5
Couple extended family	0.6020	4 331	2.5	49.2	0.4758	1 658	4.7	37.5

Notes: Results for the subgroup other income is not reported due to too few observations for computing the Gini coefficient. The adult equivalent scale where Head of household (HH) = 1.0, All other adults = 0.7, Children under 15 years = 0.5, is used when calculating this table. The exchange rate in December 1993 was 3.65 N$/US$.

Table E2 The decomposition of total inequality for consumption by subgroups, 1993/94

Subgroups	Central/southern region				North/north-east region			
	T	Con. (%)	L	Pop. (%)	T	Con. (%)	L	Pop. (%)
Urban/rural residence								
Urban	0.6841	76.8	0.6859	57.2	0.6506	21.8	0.6288	12.1
Rural	1.0940	23.2	0.8716	42.8	0.4479	78.2	0.3684	87.9
Within-group inequality	0.7653		0.7870		0.3998		0.5018	
Between-group inequality	0.0740		0.0678		0.0633		0.0775	
Main source of income								
Wages	0.6614	72.0	0.6806	70.5	0.6347	33.1	0.5991	18.9
Business	1.0529	16.8	1.2633	7.7	0.4687	4.3	0.4234	3.9
Subsistence farming	1.2462	3.1	0.8518	5.1	0.3246	52.4	0.2979	62.1
Pension	1.0523	5.9	0.9596	11.1	0.5414	8.0	0.3764	11.7
Cash remittances	0.6324	2.2	0.5967	5.4	0.3063	2.2	0.2501	3.3
Other income	0.3531	0	0.4752	0.2	0.0254	0	0.0253	0.1
Within-group inequality	0.7604		0.7714		0.3672		0.4688	
Between-group inequality	0.0789		0.0834		0.0960		0.1104	
Education of HH								
No formal	0.4538	7.1	0.4174	24.6	0.2924	25.8	0.2741	35.0
Primary	0.4370	11.1	0.3980	24.2	0.4294	34.2	0.3581	40.2
Secondary	0.6834	55.9	0.6367	42.5	0.6494	32.0	0.5553	22.0
Tertiary	0.3475	25.9	0.4084	8.7	0.4648	8.0	0.4670	2.8
Within-group inequality	0.5051		0.5518		0.3750		0.4781	
Between-group inequality	0.3341		0.3030		0.0882		0.1012	
Sex of HH								
Female	0.7909	18.8	0.7674	27.8	0.3701	39.7	0.3240	47.4
Male	0.8551	81.2	0.8477	72.2	0.6984	60.3	0.5661	52.6
Within-group inequality	0.8253		0.8417		0.4514		0.5677	
Between-group inequality	0.0139		0.0130		0.0118		0.0116	
Age of HH								
Below 25	0.4571	3.3	0.4219	7.2	0.3294	3.7	0.2879	4.6
25–34	0.6661	25.5	0.6597	26.4	0.5642	21.9	0.4927	18.9
35–44	0.6540	30.2	0.7079	27.1	0.5911	23.6	0.4846	20.1
45–54	1.1203	23.3	1.0617	17.3	0.6430	17.1	0.4797	17.5
55–64	1.0588	8.8	1.0037	10.9	0.7327	14.6	0.5300	14.9
65+	1.1344	9.0	1.1747	11.1	0.2764	19.2	0.2613	23.9
Within-group inequality	0.8196		0.8372		0.4296		0.5465	
Between-group inequality	0.0196		0.0176		0.0336		0.0328	

Table E2 continues The decomposition of total inequality for consumption by subgroups, 1993/94

Subgroups	Central/southern region				North/north-east region			
	T	Con. (%)	L	Pop. (%)	T	Con. (%)	L	Pop. (%)
Household composition								
Single	0.6431	8.2	0.7116	13.5	0.6472	4.0	0.6262	4.3
Single with children	0.6071	4.3	0.5956	5.9	0.3259	6.5	0.3185	9.8
Single extended family	0.5927	16.2	0.5411	26.5	0.4478	35.9	0.3672	38.3
Couple	0.9670	13.5	0.9675	8.3	0.5275	1.7	0.4367	2.4
Couple with children	0.6142	36.0	0.7570	22.3	0.6477	14.8	0.5142	14.5
Couple extended family	0.7563	21.7	0.6808	23.4	0.4498	37.2	0.3925	30.8
Within-group inequality	0.6838		0.7030		0.4042		0.4990	
Between-group inequality	0.1555		0.1518		0.0589		0.0803	
Total Inequality	**0.8393**		**0.8548**		**0.4631**		**0.5793**	

Notes: The adult equivalent scale where Head of household (HH) = 1.0, All other adults = 0.7, Children under 15 years = 0.5, is used.

Essay II: Wage discrimination in Namibia*

1 Introduction

Gender equity has long been a subject of debate. This discussion becomes more and more relevant for the development not only in industrialised countries, but also to a greater extent in less developed countries. In the latter, women are now entering the labour market more frequently than in the past. Most of them are occupied in the informal sector of the economy, where their jobs are related to their home production. However, female involvement in other sectors of the economy is constantly growing. Because of the increasing female labour force participation, I believe that gender discrimination in the labour market is an issue of growing importance, in developing countries.

In this paper I study the wage discrimination in Namibia by using individual information, which was collected in the 1993/1994 Namibia Household Income and Expenditure Survey (NHIES). The purpose is to analyse if the earnings differences between males and females are due to endowment differences or to discrimination. Methods developed by Blinder (1973), Oaxaca (1973) and Oaxaca & Ransom (1994), which has been widely used in developed countries are employed. However, their methods have only been used in a few studies concerning developing countries in Africa (Knight & Sabot, 1991, Appleton *et al.*, 1995 and Skyt Nielsen, 2000).[1]

A positive sign of the endowment component would be expected due to the fact that males usually are better endowed than females, for example, with a higher level of education. In this study I pay particular attention to the manufacturing, services and public sector. The aggregation of the three sectors is,

* I am grateful for valuable comments from Arne Bigsten, Lennart Flood, Eugenia Kazamaki Ottersten, Erik Mellander, Ronald Oaxaca, and seminar participants at the Research Institute of Industrial Economics, at the PhD workshop in Econometrics of Labour Market Discrimination at the Centre for Labour Market and Social Research, Århus, Denmark, August 13–14, 1997, at the workshop Development Economics, Göteborg University, October 3–4, 1997, and at the AEA conference in Perpignan, France, April 2–3, 1998.

[1] For a study concerning Brazil, see Birdsall & Fox (1991).

however, also considered as well as partition into urban and rural areas. Females would be expected to be less discriminated in rural areas, because of more similarities between the jobs. Ordinary least square estimation is compared with Heckman's (1979) estimation procedure in order to answer the question whether selectivity matters in explaining part of the earnings differences.

When previous authors have assessed discrimination in the Namibian labour market the emphasis has been on racial discrimination.[2] There is no doubt that the most important determinant of labour market discrimination in Namibia is ethnicity, because of the apartheid system that existed during the South African occupation.[3] However, I argue that, irrespective of race, the aspect of gender discrimination is interesting to study, because of the labour market segregation of males, which followed from the Labour Contract System (contracting mainly males), that was implemented by the Germans, and later developed by South Africa (Moorsom, 1989). Because of this, some discrimination against females would be expected. Simon (1984) notes in his study of racial discrimination that within specific job groups, average female wages are often lower than the corresponding average wages of males.

The paper is outlined as follows. Section 2 presents the framework of wage discrimination. Section 3 describes the data. Econometric considerations are discussed in Section 4. The results of the regressions are reported in Section 5. Section 6 gives some concluding remarks.

2 Measuring wage discrimination

Discrimination studies are traditionally based on regression analysis in the form of earnings equations in which different income related characteristics, for example, education, region and gender, are used to explain wage differences. In this section I discuss the earnings equation briefly, and consider the formulas of wage discrimination used in this paper. The implementation on the Namibian data is also considered. The analysis is based on the wage differential components introduced and used by Oaxaca (1973) and Oaxaca & Ransom (1994).

[2] For example, see Labour and discrimination in Namibia (1977) where average incomes are compared.
[3] Namibia gained independence in March 1990.

2.1 The earnings equation

Consider the traditional earnings equation (Mincer, 1974)

$$\ln w_i = \beta_0 + \beta_1 s_i + \beta_2 e_i + \beta_3 e_i^2 + \beta_4 x_{1i} + ... + \beta_{4+j} x_{ji} + u_i \quad \text{i = 1, …, n} \quad (1)$$

where ln w_i is the natural logarithm of earnings or wages for the *i*th individual, s_i denotes schooling, e_i is a measure of labour market experience, x_{ji} are other factors influencing earnings such as gender, geographical region and marital status for the *i*th individual, and u_i is a random disturbance term. The latter is assumed to be normally distributed with zero mean and constant variance. The quadratic experience term indicates that earnings are in general not constant after leaving school, but follow an inverted U-shaped curve, i.e. increase at a decreasing rate.

When analysing wage discrimination it is common to compare two groups, for example, union and non-union members, white and black, males and females. In my case gender is the base, and the two earnings functions will look like

$$\ln w_m = \mathbf{x}'_m \boldsymbol{\beta}_m + u_m \quad (2)$$

$$\ln w_f = \mathbf{x}'_f \boldsymbol{\beta}_f + u_f \quad (3)$$

where ln *w* is the natural logarithms of wages, $\mathbf{x}'$ is a vector of individual characteristics, $\boldsymbol{\beta}$ is a vector of coefficients, and u is a random disturbance term. The sub-index *m* and *f* refers to males and females, respectively. For convenience, the individual sub-index *i* is excluded.

In accordance with the properties of ordinary least square (OLS) estimates, the regressions go through the sample means. The wage differences between males and females can be expressed as

$$\ln \widetilde{w}_m - \ln \widetilde{w}_f = \bar{\mathbf{x}}'_m \hat{\boldsymbol{\beta}}_m - \bar{\mathbf{x}}'_f \hat{\boldsymbol{\beta}}_f \quad (4)$$

where ln $\widetilde{w}_i$ is the logarithm of the geometric mean of wages, $\bar{\mathbf{x}}'_m$ and $\bar{\mathbf{x}}'_f$ are the vectors of the arithmetic mean of the regressors, and $\hat{\boldsymbol{\beta}}_m$ and $\hat{\boldsymbol{\beta}}_f$ are the vectors of the estimated coefficients for males and females, respectively.

Let the difference between the male and female coefficient vectors be

$$\Delta\hat{\boldsymbol{\beta}} \equiv \hat{\boldsymbol{\beta}}_{\mathrm{m}} - \hat{\boldsymbol{\beta}}_{\mathrm{f}} \tag{5}$$

implying

$$\hat{\boldsymbol{\beta}}_{\mathrm{f}} \equiv \hat{\boldsymbol{\beta}}_{\mathrm{m}} - \Delta\hat{\boldsymbol{\beta}} \tag{6}$$

Substitution of (5) and (6) into equation (4) yields the male wage structure. In the absence of discrimination the male wages will also be applicable to the females. Hence,

$$\ln \widetilde{w}_{\mathrm{m}} - \ln \widetilde{w}_{\mathrm{f}} = \overline{\mathbf{x}}'_{\mathrm{f}}(\hat{\boldsymbol{\beta}}_{\mathrm{m}} - \hat{\boldsymbol{\beta}}_{\mathrm{f}}) + (\overline{\mathbf{x}}'_{\mathrm{m}} - \overline{\mathbf{x}}'_{\mathrm{f}})\hat{\boldsymbol{\beta}}_{\mathrm{m}}. \tag{7}$$

The female wage structure is obtained in a similar way, using the fact that in the absence of discrimination the female wages will also be applicable to the males. Thus, by substituting equation (5) and $\hat{\boldsymbol{\beta}}_{\mathrm{m}} \equiv \Delta\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}_{\mathrm{f}}$ into equation (4) yields

$$\ln \widetilde{w}_{\mathrm{m}} - \ln \widetilde{w}_{\mathrm{f}} = \overline{\mathbf{x}}'_{\mathrm{m}}(\hat{\boldsymbol{\beta}}_{\mathrm{m}} - \hat{\boldsymbol{\beta}}_{\mathrm{f}}) + (\overline{\mathbf{x}}'_{\mathrm{m}} - \overline{\mathbf{x}}'_{\mathrm{f}})\hat{\boldsymbol{\beta}}_{\mathrm{f}}. \tag{8}$$

The first term on the right-hand side of both equation (7) and (8) refers to differences in the returns that males and females receive for the same endowment of wage generating characteristics. The second term refers to differences in the endowments of wage generating characteristics, evaluated at the returns of the beta coefficients. Equation (7) and (8) are used as instruments when estimating wage discrimination, which I will now turn to.

2.2 A decomposition of wage discrimination

Next I assess a decomposition of the total wage discrimination. The males are here taken to be the advantaged group, while the females are the disadvantaged group.

A labour market discrimination coefficient, D, was introduced by Oaxaca (1973):

$$D = \frac{(w_m / w_f) - (w_m^* / w_f^*)}{w_m^* / w_f^*} \tag{9}$$

where $w_m / w_f =$ the actual male-female wage ratio, $w_m^* / w_f^* =$ the male-female wage ratio in the absence of discrimination. Taking the natural logarithms of equation (9) I obtain

$$\ln(D+1) = \ln(w_m / w_f) - \ln(w_m^* / w_f^*). \tag{10}$$

The coefficient of discrimination in equation (10) shows the relative wage effects of labour market discrimination. It does not measure how much of the differential refers to overvaluation of the males and how much refers to undervaluation of females. Hypothetically, a further division of the discrimination coefficient into two components is possible.[4] One refers to male overvaluation, i.e. the favouritism component, and the other refers to female undervaluation, i.e. the pure discrimination component. Equation (10) can be extended to

$$\begin{aligned} \ln(D+1) &= \ln(w_m / w_m^*) + \ln(w_f^* / w_f) \\ &= \ln(\delta_{m*} + 1) + \ln(\delta_{*f} + 1) \end{aligned} \tag{11}$$

where

$\delta_{m*} \equiv (w_m / w_m^*) - 1$ measures the relative differential between the actual male wages and the wages that males would have received in the absence of discrimination, and

$\delta_{*f} \equiv (w_f^* / w_f) - 1$ measures the relative differential between the wages that females would have received in the absence of discrimination and the actual female wages.

It is possible that the male-female wage differential is not entirely due to favouritism and/or (pure) discrimination. To some extent, the differential can reflect productivity differences. The total or gross male-female wage differential can be decomposed into a favouritism, a pure discrimination and a pure

[4] Originally, Cotton (1988) proposed this division.

productivity component as follows. In equation (10) the second term on the right-hand side is the male-female wage differential in the absence of labour market discrimination reflecting pure productivity differences, Q;

$$Q = (w_m^* / w_f^*) - 1. \tag{12}$$

Taking natural logarithms of equation (12) I obtain

$$\ln(Q+1) = \ln(w_m^* / w_f^*). \tag{13}$$

Substitution of (13) into (10) yields

$$\ln(w_m / w_f) = \ln(D+1) + \ln(Q+1). \tag{14}$$

The sum of the right-hand side terms is the gross male-female wage differential, G, so (14) can be written as

$$\ln(G+1) = \ln(D+1) + \ln(Q+1) \tag{15}$$

or (11)

$$\ln(G+1) = \ln(\delta_{m*}+1) + \ln(\delta_{*f}+1) + \ln(Q+1). \tag{16}$$

To implement equation (15) and (16), recall equation (7), which is the male wage structure. The left-hand side of equation (7) is equal to the gross male-female wage differential, i.e. ln(G+1). Further, assume that the male wage structure applies to both males and females in the absence of labour market discrimination. Then

$$\ln(D+1) = \bar{\mathbf{x}}_f' (\hat{\boldsymbol{\beta}}_m - \hat{\boldsymbol{\beta}}_f) \tag{17}$$

and

$$\ln(Q+1) = (\bar{\mathbf{x}}_m' - \bar{\mathbf{x}}_f')\hat{\boldsymbol{\beta}}_m. \tag{18}$$

Equation (17) shows the estimated differences of the coefficients and equation (18) shows the estimated differences in individual characteristics. Thus, the implementation of equation (15) only requires an application of equation (7).

To implement equation (16) I have to introduce the non-discriminatory wage structure, $\hat{\boldsymbol{\beta}}$, which is simply the pooled wage structure.[5] Equation (16) can now be expressed as

$$\ln(G+1) = \bar{\mathbf{x}}'_{m}(\hat{\boldsymbol{\beta}}_{m} - \hat{\boldsymbol{\beta}}) + \bar{\mathbf{x}}'_{f}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{f}) + (\bar{\mathbf{x}}'_{m} - \bar{\mathbf{x}}'_{f})\hat{\boldsymbol{\beta}} \qquad (19)$$

where the first term on the right-hand side refers to the estimate of the wage advantage for males, the second refers to the estimate of wage disadvantage for females, and the third refers to the productivity differential estimate.

2.3 The Namibian implementation

Above the theory of earnings equation and wage discrimination were described. Next, I implement it to the Namibian case. I include my explanatory variables into the earnings equation in accordance with the theory, except for the experience variable. Direct information on experience is seldom available and therefore a measure of potential experience is often used. This measure is estimated as the individual's age minus the individual's number of years of completed education minus six, where six is supposed to be the individual's age in his/her last pre-school year. The survey does not contain information on neither the number of years of education nor the final pre-school year. For this reason, age and age squared are used as proxies for experience. These measures underestimate both the first-order and the second-order effects of experience.

Because of the heterogeneity of the labour market in less developed countries I choose to analyse wage discrimination by sector, but as a complement an analyse of the discrimination for the aggregation of the three sectors are also carried out. I pay attention to manufacturing, services and the public sector. The agricultural sector, the fishing industry, people involved in private households and people working in territorial organisations are not included in the analysis. This is because in the two former categories people work basically on a casual wage employment and for the latter category there are too few observations in the sample. Further, individuals in the agricultural sector are in some form of self-employment. They do not necessarily generate cash income. Individuals working in private households mostly receive their incomes in kind.

[5] This is in accordance to the pooled structure proposed by Oaxaca & Ransom (1994).

The income structure in the urban areas is usually different from the rural income structure. Therefore these two economies are separated. But a dummy for the urban/rural residence had to be used in the manufacturing and services sector due to too few observations for females in the rural areas.[6]

3 Data

Individual data from the 1993/1994 Namibia Household Income and Expenditure Survey (NHIES) are used. The survey was conducted by the Central Statistics Office (CSO)/National Planning Commission, Windhoek, from November 1993 to October 1994. The data comprise 24 984 individuals. For my purposes the sample is reduced because only individuals of age 15 years and above and the part of the labour force are included. Individuals with missing values are omitted. As already mentioned in the previous section, I also exclude individuals working with agriculture, fishing activities, activities in private household and those employed in territorial organisation. This reduces the data to 2 703 individuals: 1 633 males and 1 070 females. Almost one third of the individuals are not reporting any wages. Thus 1 890 individuals, 1 208 males and 682 females, report a positive wage.

The individuals reported two types of wage measures. The first wage measure is a monthly wage referring to the survey month. The second wage measure is an annual wage, which was estimated by the individuals for the last twelve months, including the survey month. Further, a third wage measure can be estimated, the hourly wage. The estimates can be computed by means of available information about the number of hours worked during the last week.[7] However, the measures of hourly wages can be associated with measurement errors to a considerable extent, since working hours in Namibia are characterised by seasonal variations and day-to-day fluctuations. This is particularly the case for the unskilled and semi-skilled workers in the fishing and fish processing industries and in the agricultural industry. Since I have disregarded

[6] It would have been fruitful to distinguish between the formal and the informal sector, particularly in the manufacturing sector and in the wholesale and retail trade sector (included in the service sector). Since no such data are available, this is not possible.

[7] The number of hours worked during the week were recorded in intervals, which forced me to take the mid-point in each interval except for the interval, 40 hours or more, where 40 is used. To arrive at hours worked per month I multiplied weekly mid-point hours worked by 4.3.

these industries the hourly wage probably becomes more reliable. While I am aware of the fact that the model by Mincer (1974), which is used, has been derived in terms of hourly wages, I will compare the three types of wage measures, hourly, monthly and annual wages. The reason for comparing these three measures is that the monthly wage would be the measure with least measurement errors.[8] Distribution of monthly wages among females and males for the whole sample is shown in *Figure A1* in Appendix A. Most of the individuals have a monthly wage between N$ 200 up to N$ 4 000.

The independent variables used for estimating wage discrimination are age, age square, education, region, urban/rural residence, marital status and children. Nationality, being a Namibian or not, is available but there were only about two per cent of the adjusted sample that were not Namibians.[9] Further, the few outliers that were found are included, since they had a minor impact on the estimated wage differential components. A more detailed description of the variables is presented below. In addition, the three sectors, which I distinguish between, are described.

Level of educational attainment measures the highest level attained by the individual. Four levels, which also are recorded in the CSO survey, are distinguished (CSO, 1996a). The first level is no formal education, comprising individuals with no schooling or primary school Grade 1. The second level, primary education, includes individuals in Grade 1–6. The third level, secondary education, comprises individuals in Grade 7–12, where Grade 7–9 refers to junior secondary high school and Grade 10–12 refers to senior secondary high school. The fourth level tertiary education refers to individuals above Grade 12.[10] No formal education is the reference group. *Figure A*2 in Appendix A shows the level of educational attainment among females and males for the whole sample. The figure suggests that most of the females and the males for the whole sample have secondary education. Females seem not to be less educated than their counterparts.

[8] Wages are only measured as cash wages, and the wages in kind that an individual might have received are not considered.

[9] Including this variable did not have any important impact on the measurement of wage discrimination.

[10] This includes courses of study leading to a diploma or certificate issued by a university, technician or technical institute. University courses leading to first degrees, post-graduated courses leading to post graduate diploma, master degrees, doctorate degrees and teacher training.

Namibia is divided into thirteen regions. However, in this paper Namibia is divided into three regions. The first comprises the Windhoek City/Walvis Bay Centre. The second refers to Central/southern regions.[11] The North/north-east regions constitute the third region, which is the reference group.[12] The aim of this division is to capture the "urban elite" in the first region, while the other two are supposed to capture the differences between Central/southern regions and the North/north-east regions. Urban/rural residence is coded as a dummy variable being one if the individual is living in urban areas and zero otherwise.

Marital status is included as a dummy variable being one if the individual is married and zero if not married. Not married includes never married, widowed, separated and divorced individuals.

A question on the number of children was not available in the data, but by means of the variable relation to head of household the variable children could be derived. A dummy variable was created being one if the individual has one or more children, and zero otherwise.

Three sectors are distinguished. Firstly, the manufacturing, which also includes individuals working with electricity, gas, steam and hot water supply, construction, and mining and quarrying. Secondly, the service sector comprises individuals working in wholesale, hotels and restaurants, transport, storage and communications, financial intermediation, and real estate, renting and business activities. Thirdly, the public sector comprises individuals working in public administration and defence, compulsory social security, education, health and social work, and other community and social service activities. See CSO (1996b) for detail codes of the sectors.

Table B1 in Appendix B presents means and standard deviations for the variables used in the aggregation of the three regressions and for the sectors, separately. Worth noticing, is the unusually result that the mean values for the females are higher in secondary education than for the males. This is also the case for tertiary education in the public sector and in the aggregation of the sectors. In this case, however, I have to keep in mind the restriction, i.e. the individuals only are part of the labour force and have reported a positive wage.

[11] The Central/southern region includes Khomas (except Windhoek City), Erongo (except Walvis Bay Centre), Hardap, Karas, Kunene, Omaheke and Otjozondjupa.
[12] The North/north-east region includes Ohangwena, Omusati, Oshana, Oshikoto, Okavango and Caprivi.

Other available evidence supports, however, the finding that females have higher levels of education than males CSO (1995).

4 Econometric considerations

Most of the studies that consider wage decomposition employ the OLS estimator, but if selectivity bias is present OLS estimation leads to biased and inconsistent parameter estimate and this property is carried over to the components of the decomposition. In my case the selection bias concerns possible systematic differences between those individuals who have reported positive wages and those who have not reported any wages. To correct for selection bias, Heckman's (1979) two-stage estimation procedure is used.

In the presence of sample selectivity the wage differences between males and females using the male wage structure can be expressed as

$$\begin{aligned} \ln \widetilde{w}_{\mathrm{m}} - \ln \widetilde{w}_{\mathrm{f}} &= (\overline{\mathbf{x}}'_{\mathrm{m}}\hat{\boldsymbol{\beta}}_{\mathrm{m}} + \hat{\theta}_{\mathrm{m}}\overline{\hat{\lambda}}_{\mathrm{m}}) - (\overline{\mathbf{x}}'_{\mathrm{f}}\hat{\boldsymbol{\beta}}_{\mathrm{f}} + \hat{\theta}_{\mathrm{f}}\overline{\hat{\lambda}}_{\mathrm{f}}) \\ &= \overline{\mathbf{x}}'_{\mathrm{f}}(\hat{\boldsymbol{\beta}}_{\mathrm{m}} - \hat{\boldsymbol{\beta}}_{\mathrm{f}}) + (\overline{\mathbf{x}}'_{\mathrm{m}} - \overline{\mathbf{x}}'_{\mathrm{f}})\hat{\boldsymbol{\beta}}_{\mathrm{m}} + (\hat{\theta}_{\mathrm{m}}\overline{\hat{\lambda}}_{\mathrm{m}} - \hat{\theta}_{\mathrm{f}}\overline{\hat{\lambda}}_{\mathrm{f}}) \end{aligned} \qquad (20)$$

where $\ln \widetilde{w}_{\mathrm{i}}$ is the logarithm of the geometric mean of wages, $\overline{\mathbf{x}}'_{\mathrm{m}}$ and $\overline{\mathbf{x}}'_{\mathrm{f}}$ are the vectors of arithmetic means of the regressors, $\hat{\boldsymbol{\beta}}_{\mathrm{m}}$ and $\hat{\boldsymbol{\beta}}_{\mathrm{f}}$ are the vectors of the estimated coefficients. The selectivity effect is captured by $\hat{\theta}_{\mathrm{m}}$ and $\hat{\theta}_{\mathrm{f}}$ which are estimates of $\rho\sigma_{\mathrm{u}}$, ρ and σ_{u} being the correlation between the random error terms in the probit and wage equation and the standard deviation of the error term in the wage equation, respectively. $\overline{\hat{\lambda}}_{\mathrm{m}}$ and $\overline{\hat{\lambda}}_{\mathrm{f}}$ are the means of the inverse Mill's ratios for males and females, respectively. Equation (20) corresponds to equation (7) shown in Section 2.1, but here the selectivity term is added.

The question is now how to interpret the selectivity term. Reimers (1983) suggests that the selectivity term is deducted from the observed wage differential. Neuman & Oaxaca (1998) suggest a new approach of the decomposition issue. This approach decomposes the selectivity term into three components as follows

$$\hat{\theta}_{\mathrm{m}}\bar{\hat{\lambda}}_{\mathrm{m}} - \hat{\theta}_{\mathrm{f}}\bar{\hat{\lambda}}_{\mathrm{f}} = \hat{\theta}_{\mathrm{m}}(\bar{\hat{\lambda}}^{0}_{\mathrm{fp}} - \bar{\hat{\lambda}}_{\mathrm{fp}}) + \hat{\theta}_{\mathrm{m}}(\bar{\hat{\lambda}}_{\mathrm{mp}} - \bar{\hat{\lambda}}^{0}_{\mathrm{fp}}) + (\hat{\theta}_{\mathrm{m}} - \hat{\theta}_{\mathrm{f}})\bar{\hat{\lambda}}_{\mathrm{fp}} \qquad (21)$$

where $\bar{\hat{\lambda}}_{jp} = \sum_{i=1}^{N_{pj}} \hat{\lambda}_{ijp} / N_{jp}$ and $\hat{\lambda}_{ijp} = \phi(X'_{ij}\hat{\beta}_j) / \Phi(X'_{ij}\hat{\beta}_j)$ for j = m, f,

$\bar{\hat{\lambda}}^{0}_{fp} = \sum_{i=1}^{N_{pf}} \hat{\lambda}^{0}_{ifp} / N_{fp}$ and $\hat{\lambda}^{0}_{ifp} = \phi(X'_{if}\hat{\beta}_m) / \Phi(X'_{if}\hat{\beta}_m)$.

$\bar{\hat{\lambda}}^{0}_{\mathrm{fp}}$ is the female mean inverse Mills ratio if females face the same selection equation that the males face, where the sub-index *p* denotes a positive wage. The first term in equation (21) measures the effects of gender differences in the parameters of the probit selectivity equation on the male/female wage differential. The second term measures the effects of gender differences in the variables that determine whether an individual has a positive reported wage or a non-reported wage. The last term captures the gender difference in the wage response to the having of a positive wage.[13]

5 Empirical results

In this section results of the earnings equation are briefly presented, before moving on to the wage decomposition results. The investigation of the decomposition components is considered based on the results from employing OLS estimation. At the end the decomposition with Heckman's estimation procedure is also considered. These decomposition components are compared with those employed by OLS estimation using the male wage structure.

5.1 Results of the earnings equation

The regression coefficients for the three aggregated sectors by total, urban and rural, using monthly wage as the dependent variable are given in *Table C1, Table C2* and *Table C3* in Appendix C, respectively. The general pattern for the

[13] In addition to substitute (21) for the last term in (20) to obtain a decomposition of the selection effect. Neuman & Oaxaca (1998) also consider decompositions in which the three selection terms in (21) are allocated to either the discrimination or the endowment components of the wage differential. These latter alternative decompositions will not be considered here, however.

variables age and age square is as expected, i.e. the inverted U-shaped curve indicating that age (a proxy for experience) has a decreasingly positive effect on income. The educational variables also follow the expected pattern, increasing with level of education.[14] In the urban public sector females appear to have higher average rate of return for secondary education.

Living in Windhoek City/Walvis Bay Centre seems to have a positive effect on the income both for females and males. For males there is no significant effect of living in the Central/southern region when analysing the sectors, while there is a significant positive effect for females, but not in the public sector. The average positive effect of living in urban areas is considerable higher for males than for females.

Being married has a significant positive income effect for both males and females, but the coefficient is larger for males than for females.[15] This is, however, not true when analysing the rural areas, in the public sector, and the aggregation of the three sectors. Having children or not, has significant positive effect on the income for males, but is insignificant in the total and rural public sector as well as in the service sector in urban areas.[16] This variable is not statistically significant in any of the female regressions.

To test whether the parameter values associated with the female dataset are the same as those associated with the male dataset the Chow-test is applied (Green, 1993, p. 211). This test is applied to all sectors and the aggregation of the sectors. The observed F-statistics for urban areas exceed the critical values and thus I can reject the hypothesis that the parameters in the two wage equations are equal. However, for rural areas I cannot reject the hypothesis, since the observed F statistics are less than the critical values.

5.2 Results of the wage decomposition

Table 1 shows the results of the decomposition analysis for the manufacturing, services and the public sector, and the aggregation of the three sectors. The discrimination component is further divided into the favouritism (overpayment)

[14] Note that the parameter estimate of a dummy variable should be calculated as $e^{\beta} - 1$ before interpreting the percentage impact on the dependent variable. See Halvorsen & Raymond (1980) and Kennedy (1981).

[15] Dummy variables for marital status distinguishing between never married, married and widowed were also considered, but the estimated parameters were not statistically significant or only marginally significant.

[16] We also considered the number of children, but there was no change of the result.

component (δ_{m*}) and into the pure discrimination (underpayment) component (δ_{*f}). Worth noticing is that the anti-logarithms are reported, since the standard errors correspond to the anti-logarithms.

Table 1 The decomposition of the male-female wage differential by sectors

Sector and type of wage measure	Wage decomposition				
	G	D	δ_{m*}	δ_{*f}	Q
Manufacturing					
Annual wage	0.1263	0.4604	0.0529	0.3870	-0.2288
		(0.3735)	(0.1764)	(0.1613)	(0.0284)
Monthly wage	0.0521	0.3751	0.0443	0.3167	-0.2349
		(0.2758)	(0.1367)	(0.1208)	(0.0228)
Hourly wage	0.0306	0.3675	0.0435	0.3105	-0.2464
		(0.2728)	(0.1350)	(0.1207)	(0.0234)
Service sector					
Annual wage	0.2859	0.3524	0.1264	0.2007	-0.0492
		(0.2060)	(0.1094)	(0.0881)	(0.0264)
Monthly wage	0.2024	0.2873	0.1047	0.1653	-0.0660
		(0.1634)	(0.0887)	(0.0719)	(0.0219)
Hourly wage	0.1273	0.2043	0.0760	0.1192	-0.0640
		(0.1631)	(0.0928)	(0.0732)	(0.0229)
Public sector					
Annual wage	0.2011	0.2856	0.1181	0.1498	-0.0657
		(0.1333)	(0.0636)	(0.0614)	(0.0150)
Monthly wage	0.1110	0.1927	0.0815	0.1029	-0.0685
		(0.1035)	(0.0513)	(0.0494)	(0.0124)
Hourly wage	0.0633	0.1564	0.0667	0.0841	-0.0805
		(0.1127)	(0.0567)	(0.0547)	(0.0137)
All sectors					
Annual wage	0.1450	0.3180	0.1033	0.1945	-0.1313
		(0.1072)	(0.0511)	(0.0490)	(0.0115)
Monthly wage	0.0846	0.2601	0.0858	0.1605	-0.1393
		(0.0846)	(0.0413)	(0.0394)	(0.0095)
Hourly wage	0.0308	0.2122	0.0709	0.1319	-0.1496
		(0.0878)	(0.0440)	(0.0415)	(0.0100)

Notes: Standard errors in parentheses. The calculations of the standard errors have been carried out in accordance with the formulas given in Oaxaca & Ransom (1999).

The calculations in *Table 1* show that the gross unadjusted wage differential is varying between 3 and 29 per cent in favour of the males, depending on the

sector analysed and wage measure used. Comparing the different wage measures, it appears that the changes in the components of wage decomposition are larger when using the annual wage measure than when using the other two wage measures. The estimates of the decomposition components using the hourly wage measure are not statistically significant, except for the endowment component. However, when analysing the aggregation of the three sectors the decomposition estimates are significant using the hourly wage. The endowment component (Q), the market discrimination coefficient (D), and the pure discrimination component (δ_{*f}) changes more between the sectors than the favouritism component (δ_{m*}). The endowment component is, however, changing marginally between the service and public sector. The service sector has the highest gross unadjusted wage differential of the three sectors. This might not be unexpected, since the service sector comprises rather different occupational activities by the individuals. This sector may also comprise individuals being part of the informal sector. The results of this sector using the monthly wage measure suggest that the males are overpaid by about ten per cent and the females are underpaid by about 17 per cent. The female endowment advantage is about seven per cent. The market discrimination coefficient is estimated to about 29 per cent. The estimated endowment component is statistically significant at the one per cent level, whereas the pure discrimination component is statistically significant at the five per cent level and the discrimination component at the ten per cent level.

The manufacturing sector produces the highest figure of the pure discrimination component, regardless of which wage measure used. This high figure might be a result of the low participation rate of females and their occupational distribution in this sector. Note, this sector produces the lowest gross wage differential, 0.0521, since the female endowment advantage is as high as 23 per cent. Knight & Sabot (1991) analyse discrimination in the urban manufacturing sector in Tanzania. Their results for the discrimination and endowment component using the male wage structure are 0.056 and 0.232, respectively. Thus, the gross wage differential is 0.288. In my case, also using the male wage structure, the Namibian manufacturing sector has somewhat different results. By means of the male wage structure the gross wage differential is 0.0508, the discrimination component 0.3588 and the endowment component -0.3080. The Namibian results show that the males would earn about 31 per cent less than the females if both males and females were rewarded according to male prices.

This means that the females would have received about 36 per cent more than they actually received.

In the public sector all estimates – using the annual wage measure – of the decomposition are statistically significant, at least at the ten per cent level. The females are underpaid by about 15 per cent using the annual wage. This figure decreases to ten per cent using monthly wage and to eight per cent using the hourly wage. The former is statistically significant at the five per cent level.

Comparing each separately gross wage differentials with the gross wage differential for the aggregation of the three sectors only, a large difference between the service sector and the aggregation of the three sectors is found. However, comparing D, δ_{m*}, δ_{*f}, and Q the results indicate that it is fruitful to separately analyse the sectors. Including occupational dummies into the all-sector regression marginally changed the wage decomposition.[17]

Because of differences in economic activities and opportunities, or differences in levels of educational attainment, the wage structure of the urban areas may be different from the rural wage structure. Due to too small sample, a division of the manufacturing sector into urban and rural areas is not possible. For the same reason, the service rural sector is not analysed. *Table 2* shows the results of the decomposition by urban and rural areas.

[17] The inclusion of the occupational dummies reduced the effect of the educational variable, particularly for the tertiary level. See, for example, de Beyer & Knight (1989) for a discussion about the importance to include occupation into the earnings functions.

Table 2 The decomposition of the male-female wage differential by sectors and by urban and rural areas

Sector and type of wage measure	Wage decomposition				
	G	D	δ_{m*}	δ_{*f}	Q
Service sector/Urban areas					
Annual wage	0.2949	0.3824	0.1385	0.2143	-0.0632
		(0.2186)	(0.1173)	(0.0901)	(0.0254)
Monthly wage	0.1883	0.2963	0.1095	0.1684	-0.0833
		(0.1727)	(0.0951)	(0.0740)	(0.0213)
Hourly wage	0.1204	0.2206	0.0831	0.1269	-0.0821
		(0.1740)	(0.1001)	(0.0759)	(-0.0224)
Public sector/Urban areas					
Annual wage	0.2173	0.2649	0.1121	0.1374	-0.0376
		(0.1581)	(0.0777)	(0.0722)	(0.0181)
Monthly wage	0.1472	0.2020	0.0867	0.1060	-0.0456
		(0.1282)	(0.0642)	(0.0605)	(0.0155)
Hourly wage	0.0820	0.1478	0.0643	0.0785	-0.0573
		(0.1393)	(0.0710)	(0.0676)	(0.0172)
Public sector/Rural areas					
Annual wage	0.1989	0.3133	0.1248	0.1675	-0.0871
		(0.2390)	(0.1101)	(0.1116)	(0.0280)
Monthly wage	0.0762	0.1672	0.0690	0.0919	-0.0780
		(0.1746)	(0.0865)	(0.0858)	(0.0227)
Hourly wage	0.0530	0.1596	0.0660	0.0878	-0.0920
		(0.1929)	(0.0968)	(0.0940)	(0.0246)
All sectors/Urban areas					
Annual wage	0.1643	0.3341	0.1087	0.2033	-0.1273
		(0.1211)	(0.0588)	(0.0539)	(0.0124)
Monthly wage	0.0936	0.2747	0.0908	0.1686	-0.1421
		(0.0961)	(0.0477)	(0.0438)	(0.0103)
Hourly wage	0.0426	0.2184	0.0733	0.1352	-0.1443
		(0.1000)	(0.0509)	(0.0464)	(0.0111)
All sectors/Rural areas					
Annual wage	0.1014	0.2975	0.0957	0.1842	-0.1511
		(0.2346)	(0.1089)	(0.1133)	(0.0285)
Monthly wage	0.0681	0.2311	0.0757	0.1445	-0.1324
		(0.1816)	(0.0878)	(0.0888)	(0.0233)
Hourly wage	0.0222	0.1900	0.0629	0.1195	-0.1546
		(0.1857)	(0.0925)	(0.0912)	(0.0237)

Notes: Standard errors in parentheses. The calculations of the standard errors have been carried out in accordance with the formulas given in Oaxaca & Ransom (1999).

As expected the gross unadjusted wage differential is larger in the urban areas than in the rural areas. The results for the public sector in the urban areas investigating the monthly wage measure suggest that the males are overpaid by about eight per cent, while the females are underpaid by about ten. The latter is statistically significant at the ten per cent level. Moreover, the female wage endowment advantage is about five per cent with statistically significance at the one per cent level. The female wage endowment advantage is higher in the rural areas than in the urban areas. It is estimated to about eight per cent and is statistically significant at the one per cent level.

Investigating the aggregation of the three sectors divided by urban and rural areas small differences are found in the components, even in the gross wage differential particularly for the monthly and hourly wage measure. All of the estimates in *Table 2* relating to the urban areas are statistically significant except for the favouritism component when hourly wages are used. In this case the estimates suggest that the males are overpaid by seven per cent and the females are underpaid by about 14 per cent. The market discrimination component is estimated to about 22 per cent and the female endowment advantage is estimated to about 14 per cent.[18]

Next, I examine how the choice of estimation method affects the components in the decomposition of the wage differential. Specifically, the decomposition resulting from OLS estimation is compared with Heckman's (1979) two-stage estimation procedure. *Table 3* reports the results of correcting the male and female wage equations for selection effects for several partitions of the data and for three different wage measures.

[18] It may be relevant to say something about what variables are contributing to the endowment part of the wage decomposition. In this paper it is found that education accounts for the highest contribution of the wage differences. Adding more explanatory variables to the model gives a decreasing effect of the educational estimates (Ekström, 1997).

Table 3 The significance of lambda in the male and female regressions in manufacturing, services, public and all sectors

Sector and type of wage measure	Urban and rural areas		Urban areas	
	Lambda in male equation	Lambda in female equation	Lambda in male equation	Lambda in female equation
Manufacturing				
Annual wage	5% level	Not significant	-	-
Monthly wage	Not significant	Not significant	-	-
Hourly wage	Not significant	Not significant	-	-
Service sector				
Annual wage	Not significant	5% level	Not significant	5% level
Monthly wage	Not significant	1% level	Not significant	5% level
Hourly wage	Not significant	Not significant	Not significant	10% level
Public sector				
Annual wage	Not significant	Not significant	5% level	Not significant
Monthly wage	10% level	Not significant	1% level	Not significant
Hourly wage	Not significant	Not significant	Not significant	Not significant
All sectors				
Annual wage	Not significant	1% level	Not significant	5% level
Monthly wage	Not significant	1% level	Not significant	5% level
Hourly wage	Not significant	1% level	Not significant	1% level

The table shows that the significant selection effects are not occurring at the same time in the male's and the female's equations. Selectivity bias seems to be far more important for females than for males, particularly in the service sector, but also in the case of all sectors. When urban areas are considered rather similar results are obtained.[19] However, in this case selectivity bias seems to be important for the males in the public sector.

Table 4 compares the results of the wage decomposition components employing OLS estimation and Heckman's two-stage estimation procedure. In line with the results in *Table 3* for urban and rural areas, only the case of all sectors is considered. The results for the male and female selection, the selection corrected wage and OLS regressions are presented in *Table C4 and Tables C5–C7* in Appendix C, respectively.

[19] Rural areas are not considered since I cannot reject the hypothesis of that the parameters in the two wage equations are equal.

Table 4 Comparison of the decomposition components by employing OLS estimation and Heckman's two-stage estimation procedure, all sectors

Components of the decomposition	Annual wage		Monthly wage		Hourly wage	
	OLS	Heckman's estimator	OLS	Heckman's estimator	OLS	Heckman's estimator
ln(G+1)	0.1354	0.1354	0.0812	0.0812	0.0304	0.0304
ln(D+1)	0.3007	0.1460	0.2390	0.0100	0.1910	-0.0563
ln(Q+1)	-0.1653	-0.1687	-0.1578	-0.1563	-0.1606	-0.1598
Selectivity	0	0.1581	0	0.2275	0	0.2465
S1		0.0130		-0.0054		-0.0030
S2		-0.0043		0.0018		0.0010
S3		0.1494		0.2312		0.2485

Note: $S1 = \hat{\theta}_m(\bar{\hat{\lambda}}^0_{fp} - \bar{\hat{\lambda}}_{fp}), S2 = \hat{\theta}_m(\bar{\hat{\lambda}}_{mp} - \bar{\hat{\lambda}}^0_{fp}), S3 = (\hat{\theta}_m - \hat{\theta}_f)\bar{\hat{\lambda}}_{fp}$.

The magnitude of the selectivity term decreases if annual wages are used compared with if hourly wages are used as the dependent variable. The main part of the selection effect turns out to be the result of gender differences in the coefficients of the probit equation (S3). Both methods yield similar results for the endowment component [ln(Q+1)]. Thus, accounting for selectivity does not change the results with respect to this component. However, the discrimination component [ln(D+1)] changes quite dramatically when selection effects are accounted for. Comparing the discrimination components from the two methods, I find that the selectivity term has captured the largest part of the discrimination component. In the case where hourly wages are used as the dependent variable the discrimination components even become negative. This means, that allowing for selectivity females seems to be positively discriminated. This comparison suggests that it is important to account for selectivity bias in those cases where a significant lambda coefficient is obtained. However, a closer investigation of this issue has to be carried out, in order to explain why selection effects do not seem to matter for males.

6 Concluding remarks

This paper has addressed gender based labour market discrimination in three sectors, manufacturing, services and the public sector. The aggregation of the

three sectors was also considered. I have decomposed the earnings differences into components due to endowment and discrimination differences. The latter component was further divided into a favouritism component and a pure discrimination component when using the OLS estimation. The male wage structure is, however, used when the OLS results of the decompositions estimation are compared with the results from the Heckman's estimation procedure.

In each of the sectors three different wage measures were applied as the dependent variable. These are annual, monthly and hourly wages. The results illustrate that different wage measures yield rather different estimates. The monthly and the hourly wage measures produce the lowest wage differentials. The main differences lie in the favouritism component and the pure discrimination component. The endowment component is fairly constant.

A most interesting finding is that the Namibian females are better endowed than the males, a result manifested in a statistically significant negative sign on the endowment component. In part, higher levels of education for the females explain the negative sign. With respect to discrimination, the OLS-based results of the analysis of the differences in gross wages between males and females do suggest that females are discriminated in Namibia. Indeed, the fact that the females are better endowed than the males makes the estimated discrimination larger than the gross wage differential. The discrimination effect is reduced when the three sectors are considered separately. It appears that female workers are less discriminated in the public sector than in the two other sectors. The results also indicate that the gross wage differential might be higher in urban areas than in the rural areas, because of the fact that females are more productive in the rural areas compared with the urban.

Comparing Heckman's (1979) two-stage procedure with OLS-based results I find that accounting for selection does not affect the endowment component, but do indeed affect the discrimination component. The magnitude depends on which wage measure is used. Using the monthly wage the discrimination components is reduced by about 95 per cent, whereas using annual wage the selectivity term absorbs about half of the discrimination components. The result indicates the fact that selection bias seems to explain a large part of the discrimination, at least in the Namibian case.

References

Appleton *et al.* (1995), *Gender Differences in the Returns to Schooling in Three African Countries*, Oxford: Centre for the Study of African Economies, mimeo.

Birdsall, N & M L Fox (1991) "Why Males Earn More: Location and Training of Brazilian Schoolteachers", in ed. Birdsall N & R Sabot *Unfair Advantage, Labor Market Discrimination in Developing Countries*, World Bank regional and sectoral studies, Washington D.C.

Blinder, A S (1973) "Wage Discrimination: Reduced Form and Structural Estimates, *Journal of Human Resources*, vol. 8, no. 4, pp. 436–455.

CSO (1995) "The distribution of economic resources in the population of Namibia, Some highlights", National Planning Commission, Windhoek.

CSO (1996a) "Living Conditions in Namibia, Basic description with highlights", National Planning Commission, Windhoek.

CSO (1996b) "The 1993/1994 Namibia Household Income and Expenditure Survey NHIES, Administrative and technical report, Annex 2, The NHIES Manuals", National Planning Commission, Windhoek.

Cotton, J (1988) "On the decomposition of wage differentials", *Review of Economics and Statistics*, vol. 70, no. 2, pp. 236–243.

De Beyer J & J B Knight (1989) "The Role of Occupation in the Determination of Wages", *Oxford Economic Papers*, vol. 41, no. 3, pp. 595–618.

Ekström, E (1997) "Labour Market Discrimination in Namibia", Stockholm: The Research Institute of Industrial Economics and Göteborg: Department of Economics, Göteborg University, mimeo.

Green, W H (1993) *Econometric Analysis*, Prentice Hall, Englewood Cliffs, Second Edition.

Halvorsen, R & R Palmquist (1980) "The Interpretation of Dummy variables in Semilogarithmic Equations", *American Economic Review*, vol. 70, no. 3, pp. 474–475.

Heckman, J J (1979) "Sample selection bias as a specification error", *Econometrica*, vol. 47, no. 1, pp. 153–161.

Kennedy, P E (1981) "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations", *American Economic Review*, vol. 71, no. 4 p. 801.

Knight J B & R H Sabot (1991) "Labor market Discrimination in a Poor Urban Economy", in Birdsall N & R Sabot (eds), *Unfair Advantage, Labor Market Discrimination in Developing Countries*, World Bank regional and sectoral studies, Washington D.C.

Labour and discrimination in Namibia (1977) International Labour Office, Geneva.

Mincer, J (1974) *Schooling, experience, and earnings*, Columbia University Press, New York and London.

Moorsom, R J B (1989) "The Formation of the Contract Labour System in Namibia, 1990–1926", *Forced Labour and Migration Patterns of Movement within Africa*, Hans Zell Publishers, London, pp. 55–108.

Neuman, S & R L Oaxaca (1998) "Estimating Labor Market Discrimination with Selectivity Corrected Wage Equations: Methodological Considerations and An Illustration from Israel", Ramat-Gan: Department of Economics, Bar-Ilan University and Tucson: Department of Economics, University of Arizona, mimeo.

Oaxaca, R (1973) "Male-female wage differentials in urban labor markets", *International Economic Review*, vol. 14, no. 3, pp. 693–709.

Oaxaca, R L & M R Ransom (1994) "On discrimination and the decomposition of wage differentials", *Journal of Econometrics*, vol. 61, no. 1, pp. 5–21.

Oaxaca, R L & M Ransom (1999), "Calculation of Approximate Variances for Wage Decomposition Differentials", *Journal of Economic and Social Measurement*, vol. 24, no. 1, pp. 55–61.

Reimers, C W (1983) "Labor Market Discrimination against Hispanic and Black Men", *Review of Economics and Statistics*, vol. 65, no. 4, pp. 570–579.

Simon, D (1984) "Urban Poverty, Informal Sector Activity and Inter-Sectoral Linkages: Evidence from Windhoek, Namibia", *Development and Change*, vol. 15, pp. 551–576.

Skyt Nielsen, H (2000) "Wage discrimination in Zamibia: an extension of the Oaxaca-Blinder decomposition", *Applied Economics Letters*, no. 7, pp. 405–408.

White, H (1980) "A Heteroscedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroscedasticity", *Econometrica*, vol. 48, no. 4, pp. 817–838.

World Bank (1996) *World Development Report 1996*, Oxford University Press, New York.

Appendix A – Figures

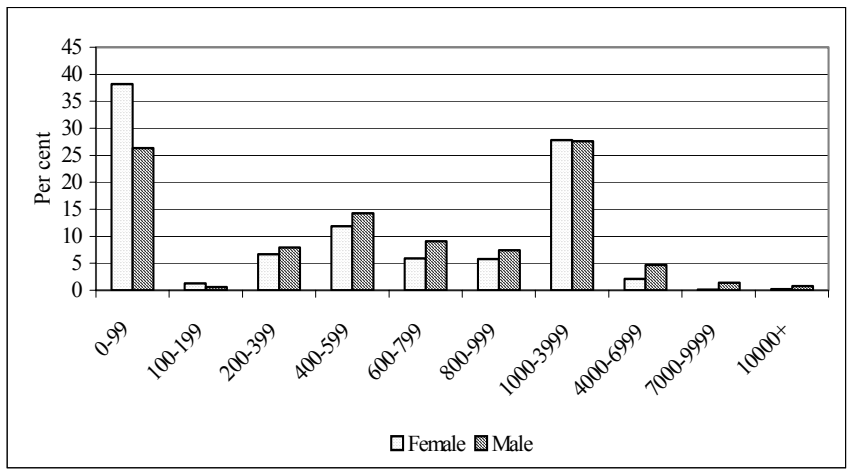


Figure A1 Distribution of monthly wages (N$) among female and males for the whole sample

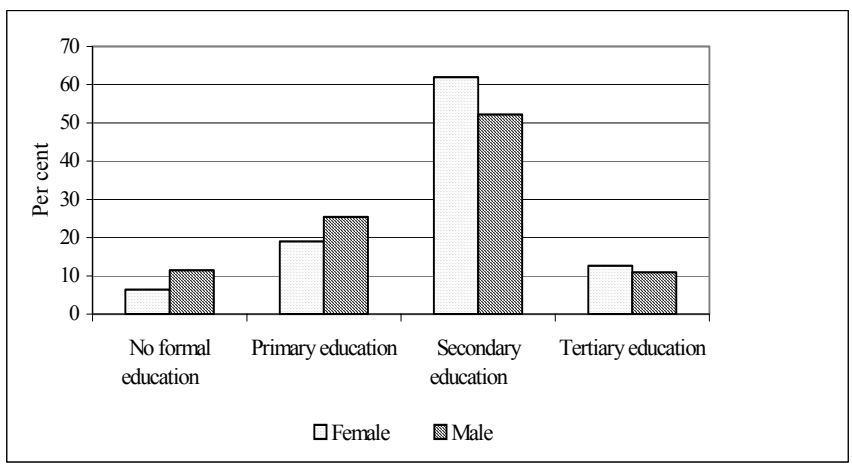


Figure A2 Highest level of educational attainment among females and males for the whole sample

Appendix B – Descriptive statistics of the variables

Table B1 Means of the variables

	Manufacturing		Service sector		Public sector		All sectors	
	Female	Male	Female	Male	Female	Male	Female	Male
Annual wage (ln)	9.054	9.173	9.126	9.378	9.298	9.481	9.218	9.354
	(8.231)	(7.810)	(7.291)	(7.482)	(6.078)	(7.018)	(6.760)	(7.466)
Montly wage (ln)	6.758	6.808	6.795	6.979	6.912	7.017	6.859	6.94
	(6.781)	(6.755)	(6.439)	(7.062)	(5.432)	(6.316)	(5.937)	(6.710)
Hourly wage (ln)	1.743	1.773	1.796	1.915	1.941	2.002	1.874	1.905
	(6.589)	(6.998)	(6.333)	(7.066)	(5.840)	(6.707)	(6.102)	(6.936)
Age	33.963	36.419	31.583	36.296	35.865	38.914	34.261	37.377
	(72.433)	(74.617)	(66.553)	(76.218)	(61.426)	(81.196)	(65.725)	(78.142)
Age squared	1259.69	1434.05	1087.72	1434.91	1359.23	1638.69	1259.45	1515.7
	(5551.48)	(6016.69)	(4958.24)	(6397.74)	(4740.19)	(6927.91)	(4968.76)	(6527.46)
Primary education	0.125	0.321	0.049	0.213	0.129	0.180	0.102	0.234
Secondary education	0.813	0.513	0.894	0.592	0.576	0.508	0.704	0.534
Tertiary education	0.049	0.039	0.047	0.090	0.257	0.226	0.168	0.128
Urban	0.847	0.761	0.903	0.880	0.630	0.610	0.741	0.735
Windhoek City/Walvis Bay Centre	0.118	0.038	0.255	0.175	0.097	0.088	0.152	0.097
Central/southern region	0.776	0.801	0.606	0.676	0.580	0.561	0.606	0.670
Married	0.630	0.681	0.480	0.692	0.558	0.757	0.538	0.715
Children	0.591	0.408	0.520	0.466	0.678	0.558	0.617	0.484

Note: Standard errors in parentheses.

Appendix C – Parameter estimates

Table C1 OLS estimates for the aggregation of the three sectors, total

Variables	Total		
	Pooled	Female	Male
Age	0.0617***	0.0706***	0.0574***
	(0.0086)	(0.0164)	(0.0106)
Age squared	-0.0006***	-0.0007***	-0.0006***
	(0.0001)	(0.0002)	(0.0001)
Primary education	0.1777***	0.0848	0.2131***
	(0.0662)	(0.1673)	(0.0722)
Secondary education	0.7543***	0.7629***	0.7992***
	(0.0609)	(0.1550)	(0.0674)
Tertiary education	1.5015***	1.4127***	1.6200***
	(0.0719)	(0.1646)	(0.0847)
Urban	0.1852***	0.0686	0.2373***
	(0.0390)	(0.0632)	(0.0483)
Windhoek City/ Walvis Bay Centre	0.9257***	1.0303***	0.9148***
	(0.0610)	(0.0885)	(0.0820)
Central/southern regions	0.1902***	0.3137***	0.1243**
	(0.0404)	(0.0644)	(0.0505)
Married	0.1993***	0.1351**	0.1524***
	(0.0377)	(0.0529)	(0.0543)
Children	0.0805**	0.0268	0.1767***
	(0.0343)	(0.0549)	(0.0446)
Constant	4.3383***	4.0921***	4.4811***
	(0.1752)	(0.3307)	(0.2164)
R2 adjusted	0.4513	0.4363	0.4849
F-statistic	156.4	53.7	114.6
No. of observations	1890	682	1208

Notes: Standard errors in parentheses. Monthly wage is used as the dependent variable. *** = significant at the 1 % level, ** = significant at the 5 % level.

Table C2 OLS estimates for the aggregation of the three sectors, urban area

Variables	Urban area		
	Pooled	Female	Male
Age	0.0592***	0.0611***	0.0531***
	(0.0091)	(0.0185)	(0.0110)
Age squared	-0.0006***	-0.0006**	-0.0006***
	(0.0001)	(0.0002)	(0.0001)
Primary education	0.1596**	0.0015	0.2084**
	(0.0781)	(0.2275)	(0.0815)
Secondary education	0.6972***	0.7559***	0.7240***
	(0.0722)	(0.2111)	(0.0766)
Tertiary education	1.4426***	1.4166***	1.5255***
	(0.0862)	(0.2227)	(0.0977)
Windhoek City/ Walvis Bay Centre	0.8001***	0.9609***	0.7614***
	(0.0657)	(0.2227)	(0.0849)
Central/southern regions	0.0299	0.2526***	-0.0948
	(0.0508)	(0.0847)	(0.0614)
Married	0.2096***	0.1691***	0.1491***
	(0.0416)	(0.0606)	(0.0579)
Children	0.1151***	0.0295	0.2504***
	(0.0381)	(0.0624)	(0.0485)
Constant	4.7165***	4.3254***	5.0169***
	(0.1908)	(0.3937)	(0.2316)
R2 adjusted	0.4730	0.4495	0.5215
F-statistic	146.3	48.5	114
No. of observations	1458	524	934

Notes: Standard errors in parentheses. Monthly wage is used as the dependent variable. *** = significant at the 1 % level, ** = significant at the 5 % level.

Table C3 OLS estimates for the aggregation of the three sectors, rural area

Variables	Rural area		
	Pooled	Female	Male
Age	0.0711***	0.0781**	0.0738***
	(0.0221)	(0.0373)	(0.0279)
Age squared	-0.0008***	-0.0009*	-0.0008***
	(0.0003)	(0.0005)	(0.0003)
Primary education	0.2472*	0.1530	0.2623*
	(0.1284)	(0.2665)	(0.1508)
Secondary education	0.9070***	0.7378***	0.9893***
	(0.1174)	(0.2446)	(0.1395)
Tertiary education	1.6845***	1.3769***	1.8946***
	(0.1381)	(0.2698)	(0.1725)
Central/southern regions	0.4043***	0.3630***	0.4220***
	(0.0733)	(0.1085)	(0.0966)
Married	0.1154	0.0838	0.0352
	(0.0848)	(0.1154)	(0.1318)
Children	0.0249	0.0174	0.0751
	(0.0774)	(0.1183)	(0.1027)
Constant	4.0424***	4.0981***	3.982***
	(0.4298)	(0.7393)	(0.5373)
R2 adjusted	0.3519	0.2974	0.3827
F-statistic	30.3	9.3	22.2
No. of observations	432	158	274

Notes: Standard errors in parentheses. Monthly wage is used as the dependent variable. *** = significant at the 1 % level, ** = significant at the 5 % level.

Table C4 Estimated regression coefficients of the selection equation

Variables	Parameter estimates for females	Parameter estimates for males
Intercept	-2.0442*** (0.0669)	-1.3297*** (0.0471)
Age	0.0405*** (0.0033)	0.0527*** (0.0024)
Age squared	-0.0005*** (0.00004)	-0.0007*** (0.00003)
Primary education	0.0702** (0.0320)	-0.0499*** (0.0171)
Secondary education	0.6632*** (0.0306)	0.1160*** (0.0165)
Tertiary education	0.7516*** (0.0359)	0.4627*** (0.0241)
Urban areas	0.4762*** (0.0157)	0.3146*** (0.0123)
Windhoek City/Walvis Bay Centre	0.5760*** (0.0233)	-0.0280 (0.0198)
Central/southern regions	0.8008*** (0.0154)	0.5816*** (0.0122)
Married	-0.1171*** (0.0158)	0.3231*** (0.0142)
Widowed/Separated/Divorced	0.1952*** (0.0256)	0.1016*** (0.0418)
Children=1	0.0676*** (0.0142)	0.1223*** (0.0116)
Service sector	0.0423*** (0.0188)	-0.0995*** (0.0120)
Public sector	0.9697*** (0.0196)	0.5075 (0.0130)

Notes: Standard errors in parentheses. Level of significance: *** = 1% level, ** = 5 % level.

Table C5 OLS and Heckman's estimator for the aggregation of the three sectors, annual wages

	OLS		Heckman's estimator	
	Female	Male	Female	Male
Age	0.0983***	0.0731***	0.0733***	0.0671***
	(0.0192)	(0.0125)	(0.0205)	(0.0160)
Age squared	-0.0010***	-0.0008***	-0.0007**	-0.0007***
	(0.0003)	(0.0001)	(0.0003)	(0.0002)
Primary education	0.1343	0.2529***	0.1191	0.2576***
	(0.2049)	(0.0859)	(0.1161)	(0.0820)
Secondary education	0.7576***	0.8901***	0.6224***	0.8740***
	(0.1899)	(0.0799)	(0.1144)	(0.0820)
Tertiary education	1.4265***	1.7872***	1.1930***	1.7336***
	(0.2016)	(0.1007)	(0.1451)	(0.1083)
Urban	0.0780	0.2667***	0.0122	0.2518***
	(0.0774)	(0.0575)	(0.0857)	(0.0706)
Windhoek City/	1.0002***	0.7819***	0.9072***	0.7967***
Walvis Bay Centre	(0.1080)	(0.0970)	(0.1188)	(0.0996)
Central/	0.3158***	0.0901	0.1282	0.0485
southern regions	(0.0789)	(0.0596)	(0.1016)	(0.0829)
Married	0.1047*	0.2191***	0.1376**	0.1902***
	(0.0626)	(0.0599)	(0.0657)	(0.0722)
Lambda	x	x	-0.5461***	-0.1971
			(0.1781)	(0.2102)
Constant	5.8632***	6.5727***	6.8745***	6.8373***
	(0.4003)	(0.2571)	(0.5066)	(0.3918)
R2 adjusted	0.348	0.411	0.358	0.411
F-statistic	41.3	94.6	39.0	85.3

Notes: The White (1980) heteroscedasticity consistent variance/covariance matrix is used to obtain corrected standard errors in the selection corrected wage regression. *** = significant at the 1 % level; ** = significant at the 5 % level.

Table C6 OLS and Heckman's estimator for the aggregation of the three sectors, monthly wages

	OLS		Heckman's estimator	
	Female	Male	Female	Male
Age	0.0729***	0.0626***	0.0519***	0.0651***
	(0.0157)	(0.0105)	(0.0168)	(0.0121)
Age squared	-0.0008***	-0.0008***	-0.0005**	-0.0007***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Primary education	0.0874	0.2069***	0.0746	0.2049***
	(0.1672)	(0.0727)	(0.0861)	(0.0580)
Secondary education	0.7631***	0.8218***	0.6498***	0.8285***
	(0.1549)	(0.0676)	(0.0864)	(0.0624)
Tertiary education	1.4137***	1.6385***	1.2180***	1.6609***
	(0.1645)	(0.0851)	(0.1119)	(0.0918)
Urban	0.0672	0.2280***	0.012	0.2343***
	(0.0631)	(0.0486)	(0.0733)	(0.0608)
Windhoek City/	1.0266***	0.8834***	0.9486***	0.8772***
Walvis Bay Centre	(0.0881)	(0.0821)	(0.0983)	(0.0837)
Central/	0.3125***	0.0996**	0.1553*	0.1169
southern regions	(0.0643)	(0.0504)	(0.0876)	(0.0728)
Married	0.1417***	0.2334***	0.1692***	0.2454***
	(0.0511)	(0.0506)	(0.0543)	(0.0613)
Lambda	x	x	-0.4578***	0.0823
			(0.1461)	(0.1812)
Constant	4.0673***	4.4365***	4.915	4.3261***
	(0.3266)	(0.2174)	(0.4138)	(0.3183)
R2 adjusted	0.437	0.479	0.447	0.478
F-statistic	59.7	124.1	55.9	111.6

Notes: The White (1980) heteroscedasticity consistent variance/covariance matrix is used to obtain corrected standard errors in the selection corrected wage regression. *** = significant at the 1 % level; ** = significant at the 5 % level.

Table C7 OLS and Heckman's estimator for the aggregation of the three sectors, hourly wages

	OLS		Heckman's estimator	
	Female	Male	Female	Male
Age	0.0636***	0.0581***	0.0391**	0.0595***
	(0.0169)	(0.0113)	(0.0190)	(0.0126)
Age squared	-0.0007***	-0.0006***	-0.0004**	-0.0006***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Primary education	0.2143	0.2391***	0.1994**	0.2380***
	(0.1807)	(0.0781)	(0.0831)	(0.0623)
Secondary education	0.7817***	0.8505***	0.6491***	0.8542***
	(0.1675)	(0.0727)	(0.0885)	(0.0646)
Tertiary education	1.4218***	1.6812***	1.1928***	1.6934***
	(0.1778)	(0.0915)	(0.1156)	(0.0990)
Urban	0.0899	0.2487***	0.0254	0.2521***
	(0.0682)	(0.0522)	(0.0777)	(0.0621)
Windhoek City/	0.9706***	0.7594***	0.8794***	0.7560***
Walvis Bay Centre	(0.0953)	(0.0882)	(0.1049)	(0.0884)
Central/	0.1726**	0.0122	-0.0114	0.0217
southern regions	(0.0695)	(0.0542)	(0.0853)	(0.0732)
Married	0.0896	0.2023***	0.1218**	0.2089***
	(0.0552)	(0.0544)	(0.0578)	(0.0651)
Lambda	x	x	-0.5356***	0.0450
			(0.1477)	(0.1913)
Constant	-0.6370*	-0.4871**	0.3549	-0.5475
	(0.3530)	(0.2337)	(0.4605)	(0.3379)
R2 adjusted	0.377	0.436	0.390	0.436
F-statistic	46.8	104.8	44.5	94.3

Notes: The White (1980) heteroscedasticity consistent variance/covariance matrix is used to obtain corrected standard errors in the selection corrected wage regression. *** = significant at the 1 % level; ** = significant at the 5 % level.

Essay III: The value of a third year in upper secondary vocational education – Evidence from a piloting scheme*

1 Introduction

Vocational education and training has been the focus of the educational policy in many European countries during the 1990s (see e.g. Lasonen (1996), Bennett *et al*. (1995), Cloonan & Canning (2000), Heraty *et al.* (2000)). In Sweden a reform of the vocational education and training[1] in upper secondary education was proposed in 1986. The Swedish Employer's Confederation (SAF) and the Swedish Confederation of Trade Unions (LO) were both promoting this expansion and upgrading of the vocational education (Lundahl, 1997). However, it was not until 1991 that the reform was decided upon (Government Bill 1990/91:85). Before the reform, the vocational programmes lasted for two years and contained a mix of general education and vocational training. The academic programmes were mainly three-year programmes. The vocational programmes were extended to three years and included increased theoretical subjects as a consequence of the 1991-reform. This also gave individuals from vocational programmes eligibility to more university programmes than before.
The reform was preceded by a pilot scheme with three-year vocational programmes. The pilot scheme was carried out in the academic years 1988/89,

* I acknowledge financial support from the Institute for Labour Market Policy Evaluation (IFAU). Åsa Murray who has been the project manager has given valuable comments and contributed with her knowledge in the vocational educational area. Peter Fredriksson has given me many helpful comments. I also thank Per-Anders Edin and Roope Uusitalo for valuable comments. Comments from Lisbeth Lundahl, Björn Öckert, seminar participants at IFAU and at the workshop on the Effects of Adult Education, Uppsala, December, 17–18, 2001 on an earlier version are also gratefully acknowledged. I thank Helge Bennmarker and Anders Skarlind for great help with the data collection, Thorbjörn Wall for providing data on the pilot scheme in upper secondary school and Kerstin Johansson for providing the composition of employment and the local unemployment statistics. The usual disclaimer applies.
[1] Further on I shall abbreviate it vocational education or vocational programme.

1989/90 and 1990/91.[2] In the first academic year of the pilot scheme, 6 000 study places were planned for the new programmes. The following year another 10 000 study places were planned for and in the last academic year 11 200 study places (SOU 1992:25). The pilot scheme involved about 12 per cent of the students in vocational programmes the first year, 19 per cent the second and 21 per cent the third year.

The pilot scheme and the reform have been described in a number of previous reports regarding the implementation of the pilot scheme and the consequences of the reform.[3] My approach is instead to use the pilot scheme to examine the value of a third year in upper secondary vocational education using register data. The purpose of the paper is to estimate the effects of the reform on the career development of the students. This is made possible by the fact that conditional on a set of local characteristics the assignment of the pilot scheme can be viewed as random. I examine the effects of the pilot scheme on three outcomes: i) years (two- or three-year) of upper secondary education (including adult secondary education), ii) university enrolment, and iii) the rate of inactivity, i.e. individuals who are neither employed nor in higher education.

The paper is structured as follows. Section 2 gives a brief description of the Swedish school system. Section 3 describes the pilot scheme. The empirical framework is presented in Section 4. In Section 5 the results are presented. Section 6 concludes.

2 The Swedish school system

To facilitate the understanding of the analysis of the paper for those who are not familiar with the school system in Sweden it is essential to briefly describe the Swedish public school system, particularly the upper secondary school system.[4]

[2] The pilot scheme was tried out in 1987/88 in a small scale with 500 study places. The two-year vocational programmes were then prolonged with a one-year supplementary course, but it was not a pilot with workplace training. Therefore, the 500-pilot cannot be compared with this pilot scheme.

[3] SOU 1989:90, 1989:106, 1990:75, 1992:25, 1996:1, 1997:1, 1997:107, National Agency for Education (1999, 2000a, 2000b).

[4] For a detailed description see National Agency for Education (2002).

The school system in Sweden comprises compulsory (e.g. the nine-year compulsory education) and non-compulsory schooling (e.g. upper secondary education and adult secondary education). The nine-year compulsory education was implemented in the 1960s and in 1971 the vocational education was integrated with the upper secondary school. The 1971-reform comprised academic as well as vocational education, including a large number of special courses. After 20 years a new reform of upper secondary education was decided upon. In this reform both academic and vocational programmes were reformed, but the major change concerned the vocational programmes. They changed from two-year to three-year programmes.

Today the upper secondary school is again faced with a new reform, which was put forward to the Ministry of Education by the Committee of upper secondary school in January 2003 (SOU 2002:120). There are a number of reasons for a new reform. One of them is the need of a broader education in working life. The purpose with the new reform is mainly to give the young people a broader upper secondary education already at the beginning of this level.

All students who have completed the nine-year compulsory school are offered an upper secondary education. A student has the right to begin a programme in upper secondary school until the first calendar year the student turns 20. In today's upper secondary school there are two academic and fifteen vocational programmes. All are three-year programmes. They offer a broad general education and give eligibility to continue to higher education. But note that a particular upper secondary programme may not necessarily give eligibility to all university studies, particularly if the students from the vocational programmes do not make use of the individual choice in an optimal way (Johansson & Strandberg, 1999). Most upper secondary schools are municipal and most students attend a school in the municipality of residence. Students may, however, choose to attend a school in another municipality, e.g. if the programme desired is not offered in their own municipality.

Students at age 20 or older may attend adult secondary education.[5] Adult secondary education is meant to offer education to adults who lack the equivalent of compulsory school or upper secondary school, but it also offers post-secondary training providing the student with knowledge in a profession.

[5] There are exceptions from this rule. People below the age of 20 may be admitted if supplies admit.

Individuals with, for example, a two-year upper secondary education are able to supplement their education leading to three-year upper secondary education. Although the adult upper secondary education is comparable to the education given in the regular upper secondary school, the two educations are not identical.

3 The pilot scheme

3.1 The implementation of the pilot scheme[6]

A modernisation of the upper secondary vocational education was a fact in the 1980s. In order to ascertain what changes that should take place, a pilot scheme was carried out to acquire experience of how to change the vocational education in the future. A reason of the pilot scheme was to increase the bridge between the schools and the working life. This was achieved through extended workplace training and increased theoretical subjects.

In this section I will describe the agencies involved in the implementation of the pilot and how these agencies affected the process. The Ministry of Education delegated the responsibility of the implementation to the National Board of Education[7].

The local school boards in the municipalities applied to participate. Usually, the initiative came from the local upper secondary school after having negotiated with the local firms about their possibilities to take on students and give them extended workplace training. This was often the most difficult part of the implementation of the new three-year vocational programmes.

The applications from the local school boards were sent to the county school board.[8] They examined the applications and made recommendations to the National Board of Education if the school should participate or not. But the National Board of Education made the final decision.[9] They had three main

[6] The information in this section is mainly based on SOU 1989:106.
[7] In 1991 it was replaced by the National Agency for Education.
[8] Upper secondary school was mainly financed by the state until 1991.
[9] By looking at data from the pilot scheme in the academic year 1990/91, one can notice that the National Board of Education followed the recommendation by the county school boards with the exception of one vocational education in a municipality.

criteria for selecting the participating municipalities: i) proportionality, ii) substitution and iii) workplace training.

The idea behind proportionality was that the number of study places within each vocational programme in the pilot scheme should be in proportion to the number of study places of the corresponding regular two-year programme. For example, if the two-year metalwork programme had five per cent of the study places of all the regular vocational programmes then the same percentage should be given to the corresponding three-year industry programme. In some cases, there were some exceptions from this criterion. Thus, the programmes with fewer number of study places were over-represented. In addition, the labour market put some limits to what was possible to obtain.

The idea behind the substitution criterion was that every class in the pilot scheme should substitute a class in the corresponding two-year programme. Even from this criterion a few exceptions were made. In some municipalities it was not possible to change all the regular classes into three-year classes within the same programme, since the number of study places in the pilot scheme was limited. For example, if one municipality had six classes with a two-year metalwork programme then a three-year industry programme did not replace all these. This made it possible for students to make a choice between a two-year and a three-year programme within the same school.

The third criterion was whether the local labour market had the possibility to arrange the extended workplace training.[10] The trade association in question examined whether the municipalities had the resources to arrange the extended workplace training. The National Board of Education made no study of their own, but relied on the recommendations by the trade representatives.

The National Board of Education received applications from local school boards for 10 100 students, but only 6 000 students could participate in the first year of the pilot scheme, i.e. 1988/89. Thus, all local school boards were not permitted to participate and the dimension of participation in the selected municipalities was also restricted.

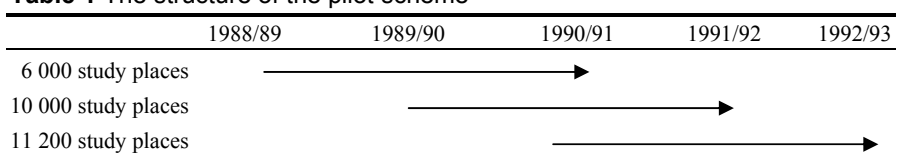
In the second year of the pilot scheme, 1989/90, it was decided that schools already participating in the pilot scheme should continue their participation by taking on a new age cohort in the first academic year of their three-year programmes. It was also decided that the number of study places in the pilot scheme should expand to include 10 000 study places.

[10] The workplace training was reduced in the 1991-reform decision.

The National Board of Education had the same criteria the second year of the pilot scheme as for the first, but another criterion was added. It concerned the extended number of study places. They should be distributed over the country to municipalities of varying size and characteristics. The recommendations by the trade associations regarding the workplace training were even more important in the second year of the pilot scheme. Altogether the local school boards applied for 15 400 students to participate. About 10 000 students could participate in the second year of the pilot scheme. Thus, a selection of local boards and the dimension of their participation were also made the second year.

The third year of the pilot scheme, 1990/91, has not been descriptively analysed. However, the number of schools in the third year was probably very much the same as the previous year, since the number of study places in the pilot scheme was only extended by 1 200 study places. The structure of the pilot scheme is illustrated in *Table 1*.

Table 1 The structure of the pilot scheme



The table shows that those individuals, who began in the academic year 1988/89, i.e. the first year of the pilot scheme, left upper secondary school in the spring of 1991.

The implementation of the pilot scheme generates a framework where students, living in municipalities that were given the pilot scheme, had the possibility of three-year vocational studies instead of the regular two-year vocational studies. The pilot scheme may be used to estimate differences in educational attainment, since attainment may be caused by the student's municipality of residence.

3.2 The distribution of the pilot scheme

At the time of the pilot scheme Sweden had 284 municipalities. But only 193 of these municipalities offered vocational education. In the first year of the pilot scheme, i.e. 1988/89, 93 municipalities and 23 counties were involved. In

the academic year 1989/90 145 municipalities and all Sweden's 24 counties had some kind of pilot scheme. Only five of these municipalities had exclusively three-year vocational education. Thus, 140 municipalities offered both three-year and two-year vocational education in some vocational fields (SOU 1990:75). Another two municipalities were added in the last year of the pilot scheme, counting to 147 municipalities. The pilot scheme seemed to be more common in larger municipalities with upper secondary education than in smaller ones.

In the academic year 1990/91 76.2 per cent of the 193 municipalities that offered vocational education were involved in the pilot scheme. A detailed table of how many pilot municipalities per county that participated in the academic year 1990/91 is presented in Appendix A.

The total number of study places in the pilot scheme were 6 236, 9 852 and 11 106 in the academic years 1988/89, 1989/90 and 1990/91, respectively (SOU 1992:25). The study places in the pilot scheme involved about 5.8 per cent of the admitted students in upper secondary school the first year. The corresponding figures for the second and the third year are 9.4 per cent and 10.5 per cent, respectively. Instead, if the comparison is made against the admitted students involved in only vocational studies the following figures are obtained: 12 per cent, 19 per cent and 21 per cent for the academic year, respectively (Statistics Sweden, Statistical Yearbook).

Before the introduction of the pilot scheme a student could chose among 17 vocational programmes in upper secondary school. Most of these were reformed into a three-year vocational programme in the pilot scheme (see Appendix B).

The pilot scheme in 1988/89 comprised ten vocational programmes. In 1989/90 seven new programmes were added. The graphic programme was included the last year of the pilot scheme. *Table 2* presents the three-year vocational programmes along with their number of study places and number of admitted students.

Table 2 Number of study places and admitted students by educational programme and year in the pilot scheme

Programme	1988/89 Number of study places	1988 Number of admitted students	1989/90 Number of study places	1989 Number of admitted students	1990/91 Number of study places	1990 Number of admitted students
Building and construction	296	287	408	403	432	429
Electrical & telecom-munication engineering	528	535	656	637	776	770
Motor and transport engineering	752	725	992	961	1 056	1 035
Commercial	210	207	660	640	990	953
Industry	1 536	1 380	1 856	1 638	1 976	1 739
Caring services	2 106	1 804	2 946	2 434	3 040	2 537
Textile & clothing manufacturing	136	122	208	163	224	186
Nursing	256	226	390	389	420	420
Heating & plumbing	64	65	72	73	104	105
Natural resource use	352	328	640	588	720	717
Craft	-	-	32	31	64	61
Restaurant	-	-	336	335	416	414
Food manufacturing	-	-	224	205	256	229
Painting	-	-	56	54	88	86
Constructional metalwork	-	-	56	51	56	52
Woodwork	-	-	144	130	168	152
Productive engineering	-	-	176	147	208	161
Graphic	-	-	-	-	112	112
Total	**6 236**	**5 679**	**9 852**	**8 879**	**11 106**	**10 158**

Source: SOU 1990:75 & Statistics Sweden, various issues.

The largest programmes in the pilot scheme were industry, caring services, and motor and transport engineering. With this background borne in mind, I will now turn to discuss the empirical framework of this study.

4 Empirical framework

4.1 Characterising the pilot scheme

After nine-years of compulsory school it is the individuals' choice to continue to upper secondary school or not. The individuals also decide what programme they will attend at upper secondary school. Therefore, individuals with

different education also differ in other respects such as interests, motivation and ability. To evaluate the effect of education on the outcome, it is necessary to disentangle the causal impact of education from outcome differentials caused by other differences between individuals. The problem is that information on these other factors usually is missing. Thus, it is difficult to identify the causal effect of education. For an excellent review of earlier studies on the causal effect of education on earnings, see Card (1999).

One approach to measure the causal effect of education is to use a situation where other variables than the individuals' own choice is relevant for how much and which type of education they obtain. In some previous studies changes in the educational system are used as an exogenous source of variation in educational attainment. For example, Harmon & Walker (1995) used the change of the minimum school-leaving age as the exogenous variation in educational attainment estimating the return to schooling for the United Kingdom. Meghir & Palme (1999) study the effect of schooling on earnings for a cohort of Swedish individuals who on one hand faced a comprehensive school system of seven- or eight-year and individuals who, on the other hand, faced a nine-year comprehensive school system at the same time. The gradual implementation of the nine-year comprehensive school system in the 1950s generated their exogenous variation in educational attainment. Both these studies forced individuals to stay longer in school.

In contrast to many previous studies, this study uses an exogenous variation in educational attainment, which did not force all individuals to longer schooling, but offered a choice of a longer schooling.[11] In the pilot scheme that preceded the 1991-reform, students were presented a choice of three-year vocational studies instead of the regular two-year vocational studies. Consider individuals who have decided to pursue vocational studies. Having a three-year vocational alternative may affect the choice margin between a vocational and academic programme as well. However, I will ignore the effects of this choice throughout the analysis. Individuals who lived in municipalities with the pilot scheme could choose three-year vocational programmes – sometimes as an alternative to two-year vocational programmes in the same field, sometimes as the only vocational programme available in a vocational field. Individuals who

[11] Denny & Holm (2000) use an educational policy reform as the exogenous variation in education, but the reform only affected part of the students – those from lower socio-economic background.

lived in municipalities with no pilot scheme could only choose among two-year vocational programmes. However, those who only had the three-year option in some fields could choose the traditional two-year programme in a nearby municipality. Also, remember that not all municipalities offer vocational education and therefore individuals living in a non-pilot municipality may live closer to a municipality that offered the three-year vocational education.

In one study (SOU 1990:75) it was found that students who had the possibility to choose between two- and three-year vocational programmes tended to choose the traditional two-year programme as their first priority. One example pertains to the caring services programme. Students preferred the two-year programme because they understood that the three-year caring services programme gave a more tight professional qualification (SOU 1990:75).

The purpose is to estimate the value of a three-year upper secondary vocational education. Why not compare those with two- and three-year vocational education? The reason why this is not done is that some pilot municipalities offered both three- and two-year vocational programmes. The students in these municipalities could choose between a range of vocational programmes of different length. Educational attainment is non-random. Instead, to be able to use the random variation in education that the pilot scheme could give rise to, the differences in educational attainment caused by living in a municipality with the pilot scheme or not are studied. Thus, I group the data based on individuals who lived in a municipality with the pilot scheme and individuals who lived in a municipality without pilot scheme, at the time of application to the upper secondary school. The two groups are denoted *pilot municipality* and *non-pilot municipality*, respectively.

What effects could the pilot scheme have had for the individuals' choice of education and their future labour market situation? The pilot scheme probably affects: i) the choice of continuing to the upper secondary school or not, ii) the choice of the programme, iii) the probability of completing the programme, iv) the choice of supplementing their education within adult secondary education or not, v) the choice of university enrolment or not, and vi) future earnings and employment. In this paper, I measure the effect of the pilot scheme on years of upper secondary education, university enrolment and the rate of inactivity. Instrumental variable estimates of the effect of years of upper secondary education on university enrolment and the rate of inactivity are also presented.

4.2 Data

To study the effects of the pilot scheme I have used data from registers on education and employment. The data cover the working age population in Sweden and the years 1990–1998. Along with these data the National Agency for Education provided information on the participating municipalities in the pilot scheme in 1990/91.[12] In addition, municipality specific variables such as the unemployment rate for the years 1990–98, the share of highly-educated, and the composition of employment for the year 1990 are available.

I decided to examine the last year of the pilot scheme, i.e. 1990/91, for two reasons. The first reason is that the pilot scheme had been running for two years. This increases the probability that students from compulsory school had sufficient information about the new programmes. The second reason is that data on grade point averages from compulsory school were available for the age cohort who left compulsory school in 1990.

A sample of individuals born in 1974 was drawn from the registers. Individuals born in 1974 normally begin upper secondary school in 1990. The first restriction was that they should have left the nine-year compulsory school at age 16 in 1990. The second restriction is that the individuals should have completed a three-year vocational programme in the pilot scheme or a regular two-year vocational programme three or two years later, i.e. in 1993 or 1992, respectively, which is also the normal rate of studies. The reason for this restriction is that I only have information from which programme the individuals left. The comparison group then consists of individuals who begin upper secondary school at the same time as the study group and finished in 1992. An alternative comparison group is individuals who completed the upper secondary education at the same time as the study group, i.e. in 1993. Since the academic year 1990/91 was the last year of the pilot scheme, individuals in the alternative comparison group are not part of the pilot scheme and may therefore differ from those within the pilot scheme. Thus, I chose to compare the study group with the former comparison group. As was mentioned earlier, I divide the data into individuals who lived in a *pilot municipality* and individuals who lived in a *non-pilot municipality* in 1990.

The total number of individuals used and the sample restrictions made in the analysis are presented in detail in Appendix C. The data consist of students in

[12] The data were generated at the time of the pilot scheme.

all vocational programmes available in the pilot scheme and the regular system. But I excluded the students in programmes, which did not have a corresponding programme either in the pilot scheme or in the regular system. This exclusion is due to that new and old programmes do not have a natural comparison group. A number of individuals were also excluded due to inadequate data. After the restrictions the data consist of 21 602 individuals in the pilot municipalities and 8 475 in the non-pilot municipalities.[13]

I have three outcome variables. They are all defined as indicator variables. The first is an indicator of three-years of upper secondary education (including adult secondary education). The second is an indicator of university enrolment. University enrolment has been measured by data from the register of students who attend higher education. This variable equals one the year the individual register at the university and in the following years, otherwise zero. Finally, I have an indicator for inactivity. An individual is defined as inactive if neither employed nor in higher education.

The individual background variables are: grade point average when leaving the compulsory school in 1990; gender; if individuals immigrated to Sweden; and county of residence in 1990. In addition to the individual variables, municipality specific variables such as the local unemployment rate, the share of highly-educated and the composition of employment in 1990 are included as controls. The municipality variables are meant to capture variations between the municipalities that were important for the determination of the pilot status. The local unemployment rate is defined as the number of individuals who can take employment at once, divided by the number of individuals in the labour force. The share of highly-educated is the sum of post secondary and tertiary education in the population. Industry composition is measured by 43 variables and is defined as the number of individuals in an industry divided by the number of individuals in the population aged 18–65.

Descriptive statistics of the variables (except the composition of employment and the county dummies) used in the regression analysis are presented in *Table 3*. The composition of the data changes as one move through the studied period. It is mainly the dependent variables that change and this is because of their definitions. Thus, descriptive statistics are reported each year for those variables that change.

[13] Individuals in a non-pilot municipality may have lived close to a pilot-municipality, and thus these individuals completed a three-year vocational programme. The reverse is also true.

Table 3 Descriptive statistics of variables used in the analysis

	1992 Mean	1993 Mean	1994 Mean	1995 Mean	1996 Mean	1997 Mean	1998 Mean
Pilot municipality							
Outcome variables							
Three years of USE		0.49	0.51	0.52			
University enrolments		0.03	0.05	0.08	0.10	0.12	0.14
Inactivity	0.18	0.18	0.14	0.09	0.08	0.08	0.06
Background variables							
Female	0.39						
Immigrant	0.05	0.04					
Grade point average	2.92 (0.49)	2.93 (0.49)					
Unemployment rate 1990 in the municipality	0.019 (0.008)						
Proportion of highly-educated 1990 in the municipality	0.152 (0.050)	0.149 (0.049)					
Number of individuals	16 097	21 584	21 564	21 533	21 478	21 414	21 323
Non-pilot municipality							
Outcome variables							
Three years of USE		0.42	0.45	0.46			
University enrolments		0.03	0.05	0.07	0.09	0.11	0.12
Inactivity	0.17	0.16	0.12	0.07	0.07	0.07	0.05
Background variables							
Female	0.38	0.40					
Immigrant	0.04						
Grade point average	2.92 (0.49)	2.94 (0.49)					
Unemployment rate 1990 in the municipality	0.014 (0.010)	0.015 (0.010)					
Proportion of highly-educated 1990 in the municipality	0.132 (0.054)	0.130 (0.053)					
Number of individuals	7 068	8 467	8 449	8 433	8 418	8 385	8 347

Notes: Standard deviations are reported in parentheses. USE = upper secondary education.

Table 3 shows that individuals, both in the pilot and non-pilot municipalities, add a third year in upper secondary education, but it comes to an end in 1995. Having a third year is more common in the pilot municipalities, since these municipalities offered the three-year vocational programmes. The increase of

university enrolment reflects the fact that the transition to university education takes some time.[14] The evolution of inactivity also confirms that individuals are more in education or employment in later years. The individual background variables are similar in the two groups, whereas the municipality specific variables are higher in the pilot municipalities.

4.2.1 Assumptions

To use the randomisation that the pilot scheme possibly involved, one has to use the population that was affected by the pilot scheme, e.g. use all individuals who leave the nine-year compulsory school a specific year. As was outlined in the data section I only have data on individuals who were born in 1974 and who left the nine-year compulsory school in 1990. Also, I do not observe whether an individual begin a two- or a three-year education. Thus, some assumptions are needed in order to interpret the results.

I have to assume that there are no systematic differences between individuals in the pilot municipalities and non-pilot municipalities with respect to i) the choice to continue to upper secondary school or not, ii) the choice of the programme, and iii) the probability of completing the programme. The plausibility of assumptions i) and ii) are checked using information about the application behaviour. The information is available at the level of resident municipality and therefore a division into pilot and non-pilot municipalities can be made. Information from the year 1986 – a year prior to the pilot – and the year 1990, which is the year used in this study are examined.

Ideally, for the first assumption one would like to have figures on those who actually begin upper secondary school. However, I use information about the qualified first hand applicants, which leave the nine-year compulsory school in the respective years. The result indicates that the number of students who apply to upper secondary school when leaving the nine-year compulsory school has not changed between 1986 and 1990, suggesting that there do not exist any systematic differences in continuing to upper secondary school between individuals living in a pilot or a non-pilot municipality.[15]

[14] Statistics Sweden (2002) reports that 10.6 per cent and 7.5 per cent continued to university within three years among those who left upper secondary school with a three- or a two-year vocational education in 1992/93 and 1991/92, respectively.
[15] The results are based on own calculations using the sources Statistics Sweden (1987, 1991).

The second assumption assumes that the three-year vocational programmes did not attract students from the academic programmes. One way to examine this is to present the number of the qualified first hand applicants to upper secondary school by type of programme. Unfortunately, there is no separate information about the three-year vocational programmes, but one can still examine whether the vocational programmes as a whole have attracted students from the academic programmes. The information is presented in *Table 4*.

Table 4 Qualified first hand applicants to upper secondary school programmes in 1986 and 1990 by pilot status

Programme	Pilot municipality		Non-pilot municipality	
	1986	1990	1986	1990
3-year academic	34.9 %	39.8 %	34.6 %	38.9 %
2-year academic	7.4 %	6.1 %	6.9 %	5.9 %
2- and 3-year vocational	46.1 %	46.1 %	47.5 %	47.6 %
Special courses	11.5 %	8.0 %	11.0 %	7.6 %
Total	**100.0 %**	**100.0 %**	**100.0 %**	**100.0 %**

Note: In 1986 there only existed 2-year vocational programmes.
Source: Own calculations using the sources Statistics Sweden (1987, 1991).

The table indicates that the application behaviour do not differ much by pilot status.

The third assumption is that the drop out rate is the same for individuals beginning a two- or a three-year vocational education. There is no available information about the drop out rate for individuals who continued to a two- or a three-year vocational education in 1990. But a reasonable approximation of the drop out rate for the two vocational paths can be calculated by taking admitted students in 1990 from the two-year and from the three-year vocational programmes. Then divide these by students leaving school in 1991/92 (the two-year students) and 1992/93 (the three-year students), respectively. Calculating the drop out rates gives six per cent for students leaving from a two-year vocational programme and 16 per cent for students leaving from a three-year vocational programme (Statistics Sweden, Statistical Yearbook for Sweden). Individuals dropping out from a three-year vocational programme may change to a regular two-year vocational programme explaining the low drop out rate for the two-year vocational programmes. Thus, the drop out rate for the two-year vocational programmes may be higher than six per cent. Although this

drop out rate is an approximation, and perhaps should be taken with caution, the difference in the drop out rate is a source of concern.

The pilot scheme was planned and conscious choices were made on how the reform was implemented in the municipalities. For evaluation one may ask whether the pilot scheme was implemented so that there were no pre-existing differences in the treatment or comparison group. To make sure that pre-existing differences were not present between young people before the implementation of the pilot scheme the analysis is repeated using data from 1986, a year prior to the pilot. This analysis uses the same municipality specific characteristics and the pilot status variable. Thus, all independent variables are the same. However, the dependent variable is not the same. The dependent variable is the rate of activity and exists for three different groups of young people: 16–19, 20–24 and 16–24 years old. The test is based on the coefficient of the pilot status dummy. If this coefficient is different from zero then pre-existing differences exists. The test indicates that there are no differences between young people in the pilot and non-pilot municipalities in the pre-period, since the pilot parameter estimate is not statistically significant.

4.3 Model specification

This section outlines the reduced form equations, which I use to estimate the effect of living in a pilot municipality in 1990 with respect to i) years (two- or three-year) of upper secondary vocational education (including adult secondary education)[16], ii) university enrolment, and iii) the rate of inactivity. I estimate the effect including all vocational programmes – no regressions are made separately for each vocational programme. The following reduced form equation is estimated separately for each year:

$$y_{it} = \gamma + \mathbf{X}_i'\beta + \beta_1 Z_i + \varepsilon_{it} \qquad (1)$$

where $\mathbf{X}_i'$ is a vector of individual and municipality specific characteristics, Z_i is a dummy variable and equals one if the individual lived in a pilot municipality in 1990, and ε_{it} is the error term. Equation (1) can be used to estimate all three outcomes mentioned above. Individuals living in a pilot

[16] Further on I shall abbreviate it years of education.

municipality in 1990 got access to a three-year vocational education beyond the two-year vocational education. This effect is picked up by the coefficient on the dummy variable, Z_i. The reduced form equations are estimated by ordinary least square (OLS).

The pilot scheme also gives me the opportunity to estimate the causal effect of an additional year at upper secondary school on university enrolment and on the rate of inactivity. Z_i is then used as an instrumental variable (IV), as the exogenous source of variation in education. The IV estimator is used when ordinary least square estimates of the returns to education are not consistent because years of education are endogenous. Since the model is exactly identified the IV estimate (β_{IV}) can be written as a ratio between the reduced form estimates of living in a pilot municipality on e.g. university enrolment and of living in a pilot municipality on years of education:

$$\beta_{IV} = \frac{E[y_{it} \mid Z_i = 1] - E[y_{it} \mid Z_i = 0]}{E[s_{it} \mid Z_i = 1] - E[s_{it} \mid Z_i = 0]} \qquad (2)$$

where y_{it} is university enrolment and s_{it} is years of education. Since Z_i is a dummy variable β_{IV} is the so-called Wald estimate (Wald, 1940), i.e. the mean differences in university enrolment for individuals in a pilot and non-pilot municipality divided by the mean differences in years of education for individuals in a pilot and non-pilot municipality.

5 Results

The outcomes that I have chosen in order to estimate the effect of living in a pilot municipality in 1990 are i) years of education, ii) university enrolment, and iii) rate of inactivity. One reason for examining the first outcome is that there is a possibility to adding a third year of education through other educational path, for example via adult upper secondary education. The second outcome is investigated, since one of the goals of the reform was to facilitate entry into higher education for school leavers from vocational education. Another aim with the third year was that the individual should be less vulnerable for unemployment, since a broader education with increased focus on

theoretical subjects should make the individual more flexible on the labour market. It seems reasonable, therefore to investigate the rate of inactivity. The estimation results of the effect of living in a pilot municipality in 1990 (*pilot* parameter) with respect to above mentioned outcomes are presented in this section.

Details on the pilot parameter estimates as well as the estimates on the background variables – gender, immigrant, grade point average when leaving compulsory school, local unemployment rate and the share of highly-educated – are presented in Appendix D, *Table D1*. The regressions also control for the composition of employment and the residence county in 1990.[17] The estimates on these two sets of controls are not reported, but are available upon request. Including the local unemployment rate and the share of highly-educated seems to be two reasonable municipality specific variables, which influence the choice of continuing to study. Since the industry structure was one of the criteria for obtaining the pilot, it also seems reasonable to include the composition of employment as an approximation of the industry structure. Two alternative specifications, beyond the presented specification, have also been tested. In the first alternative specification I add indicator variables for the county of residence when measuring outcomes. In the second alternative specification I also add the local unemployment rate at the time of the outcome. The results from these two specifications differ slightly from the baseline specification, but the differences are not statistically significant from zero.

According to the assumptions outlined in section 4.2.1 I should be concerned about the differences of the drop out rate between the two- and three-year vocational programmes. It is not clear whether the potential difference in the drop out rate implies that the effects are upward or downward biased. However, it is important to have in mind that there may exist a difference in the drop out rate between the two- and three-year vocational programmes when interpreting the results.

5.1 Effects on years of upper secondary education

Since individuals with the two-year education have the possibility of adding a third year of education in, for example, adult upper secondary school, the effects on years of education is examined. The estimates of the *pilot* parameter,

[17] In 1990 there were 24 counties, but during the 1990s the number has reduced to 21. Similar results emerge, irrespective of how many counties are used. I use the latter one.

i.e. living in a pilot municipality in 1990, from the reduced form on years of education are plotted by year in *Figure 1*. The dashed lines are the 95 per cent confidence bands.

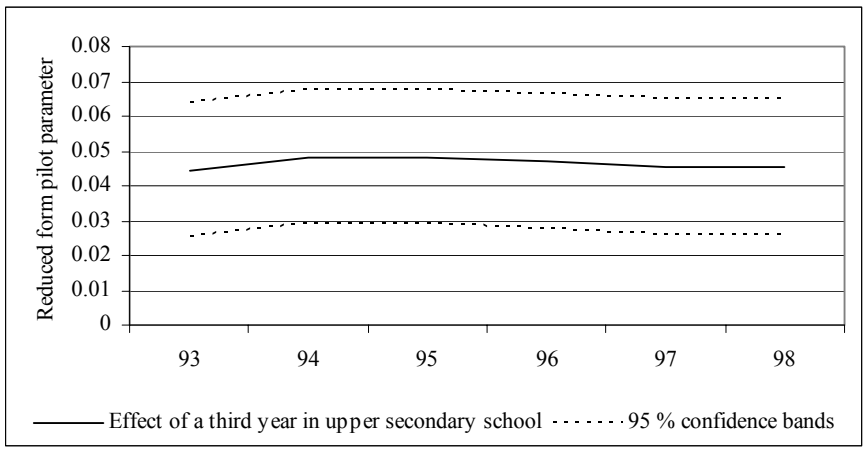


Figure 1 Reduced form estimates of the effect of living in a pilot municipality on years of education

These estimates show that individuals who lived in a pilot municipality in 1990 have a higher probability of having a three-year vocational education than individuals who lived in a non-pilot municipality. The results should be interpreted as follows: Individuals who lived in a pilot municipality in 1990 had 4.5 percentage points higher probability of having a three-year vocational education in 1993 (given that vocational education was chosen) than those living in a non-pilot municipality in 1990. The figure further indicates that if individuals with a two-year vocational education added a third year of education, they did so at the same extent in both pilot statuses, since the effect is constant over time.

5.2 University enrolment rate effects

One goal of adding a third year to vocational education was to facilitate university enrolment for graduates from vocational programmes. The third year increased eligibility for university studies. Therefore, it is of particular interest to study the university enrolment rate effect to examine whether individuals who lived in a pilot municipality enrolled at the universities to a greater extent.

Figure 2 shows the effects of living in a pilot municipality in 1990 on university enrolments by year. The 95 per cent confidence bands are the dashed lines.

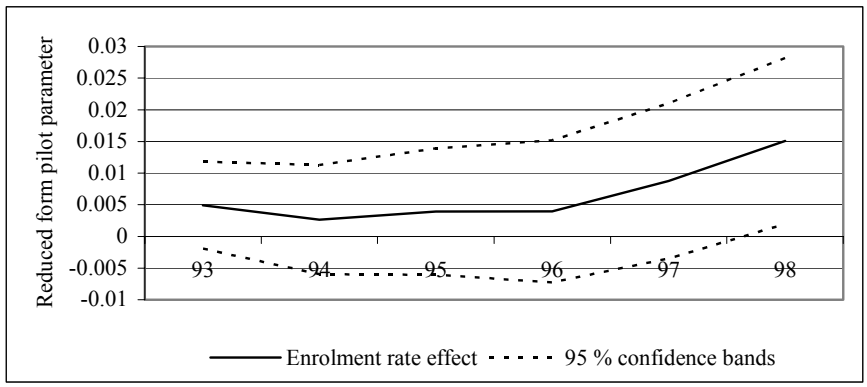


Figure 2 Reduced form estimate of the effect of living in a pilot municipality on university enrolments

There is a positive relationship between living in a pilot municipality in 1990 and the probability of university enrolment, but the effect is only statistically significant in 1998. The fact that the effect increases over time is probably due to the fact that Swedish young adults usually do not continue directly to higher education after leaving upper secondary school. The median age for enrolling at the university was 22.7 years in 2000 (OECD, 2002). Note that the effect is very small, but increases during the period. In 1998, individuals living in a pilot municipality have a 1.5 percentage point higher probability of enrolling at the university if the individual lived in a pilot municipality.

The goal to facilitate university enrolment is an interesting and important issue. This issue should be of general interest, since a broader and more general education in the upper secondary school affects the eligibility to university studies. The three-year vocational programmes offered more theoretical education, which increased the possibilities of continuing to higher education for individuals with this education compared with individuals having a regular two-year vocational education. Being eligible for more university studies should off course cause an increase in university enrolments, but what is the magnitude? I now turn to consider this effect by using the instrumental variable (IV) technique.

As was outlined in Section 4.3 the instrument, Z_i, is binary and equals one if the individual lived in a pilot municipality in 1990, zero otherwise. In this case, the IV estimates should be interpreted as the average effect of a third year of vocational education on university enrolment for the individuals whose behaviour was changed by living in a pilot municipality in 1990.[18] *Figure 3* plots the IV estimates of the effect of years of education on university enrolments by year along with the 95 per cent confidence bands.

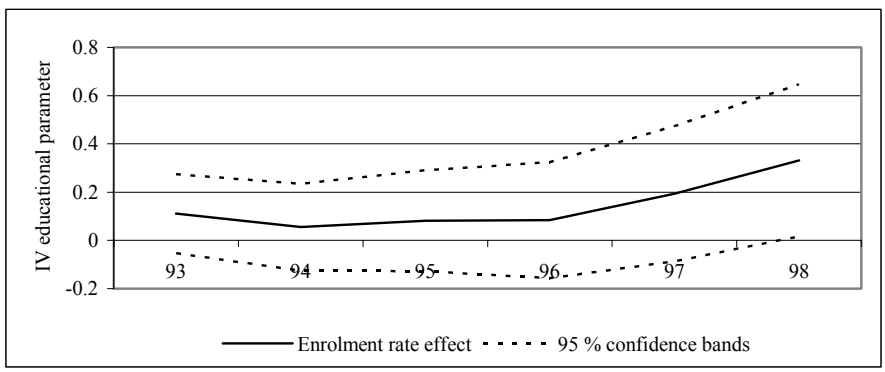


Figure 3 Instrumental variables estimates of the university enrolment rate effects

The IV estimates for the enrolment rates are positive, but the effect is only statistically significant in the last year. The interpretation of the positive university enrolment rate effects on an additional year of vocational education is that these individuals are, to a higher extent, eligible to university studies than those with a regular two-year vocational education. The probability of continuing to university within six years is about 33 percentage points higher for school leavers who had a three-year vocational education and whose educational attainment was affected by living in a pilot municipality.[19] The corresponding OLS and the probit estimates (see *Table D2* in Appendix D) are

[18] This interpretation is due to Imbens and Angrist (1994).

[19] Similar results emerge controlling for other municipality specific variables such as mean labour income in 1990 or population growth between 1989 and 1990. However, controlling for the size of the municipality in 1990 the university enrolment rate effect is 0.143 in 1998, but the effect is not statistically significant. Estimating separate regressions for females and males the university enrolment rate effect turns out to be insignificantly negative for the males.

3.3 and 3.4 percentage points, respectively. In the first year the effect is even significantly negative. By using OLS the effects are considerably under-estimated.

5.3 Inactivity rate effects

An essential aim of the 1991-reform was to make the individual more flexible and usable on the labour market through a broader education with increased focus on theoretical subjects (Gov. Bill 1987/88:102). The individual should, therefore, be less vulnerable for unemployment and exclusion from the labour market. Therefore, it would be of particular interest to study the effect of living in a pilot municipality in 1990 on unemployment. Given that the third year also increased the possibility to continue to higher education, I have chosen to analyse the rate of inactivity; an individual is defined as inactive if she is neither employed nor in higher education. *Figure 4* shows the inactivity rate effect of living in a pilot municipality in 1990 by year. The 95 per cent confidence bands are shown by the dashed lines.

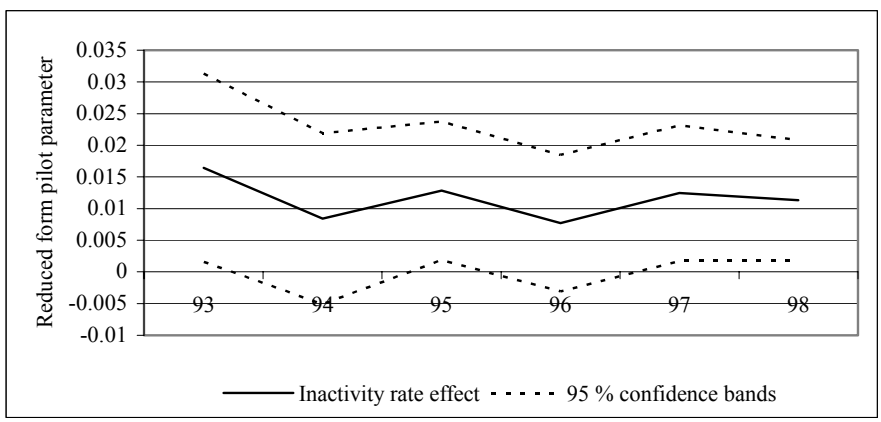


Figure 4 Reduced form estimates of the effect of living in a pilot municipality on rate of inactivity

Looking at the figure it seems that government objectives were not fulfilled. Instead the figure shows a positive effect of the inactivity rate for individuals who lived in the pilot municipalities. However, the effect is greatest in the first year of the studied period. It should be noted that unemployment in Sweden was continually very high during the 1990s until 1998 when the unemployment

rate began to decrease (Statistics Sweden, 2002). Among young adults it was even higher during this period (Edin *et al.*, 2000).

Might the choice of the comparison group influence the result? A sensitivity analysis was carried out using individuals with a two-year vocational programme completing the upper secondary education at the same time as the individuals with the three-year vocational programme, i.e. all individuals left the upper secondary school in 1993. The sensitivity analysis was carried out for the last year of the studied period. The effect of living in a pilot municipality on the rate of inactivity turns out to be zero.

The differences in the results might be explained by the changes in unemployment rate between 1992 and 1993. It should be noted that the unemployment among the young (16–24 years) was 11.4 per cent in 1992, whereas the same figure had increased to 18.4 per cent in 1993 (Statistics Sweden, 2002). The consequence of not directly be able to get into the labour market on completing the studies may give long-terming unemployment effects; see for example Gregg (2001).

What is the effect of an additional year of vocational education on inactivity? *Figure 5* plots the IV estimates of the effect of years of education on the rate of inactivity. The dashed lines are the 95 per cent confidence bands. The IV estimates should be interpreted as the average effect of a third year of education on the rate of inactivity for the individuals whose behaviour was changed by living in a pilot municipality in 1990.

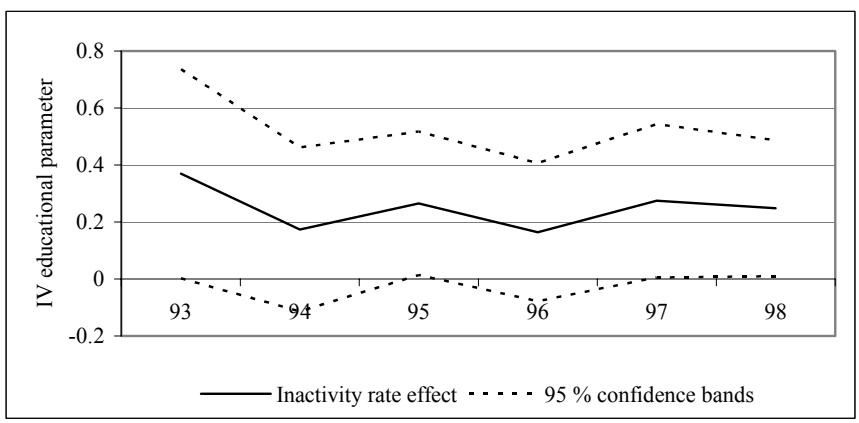


Figure 5 Instrumental variables estimates of inactivity rate effects

The probability of being inactive is positive and significantly different from zero in four of the years.[20] In 1997 and 1998, individuals with a three-year vocational education had about 25 percentage points' higher probability of being inactive relative to individuals with a two-year vocational education.[21] The positive inactivity rate effects may reflect the fact that the individuals with a two-year vocational education managed to establish themselves on the labour market before the unemployment crisis. These results differ from the OLS and probit estimates. From 1995 and onwards the effects are negative and significantly different from zero, see *Table D3* in Appendix D. Moreover, the IV estimates are significantly different from the OLS estimates for the years 1995, 1997 and 1998.[22]

What happens to the estimate with the other comparison group? The effect of a third year in vocational education is closer to zero and it is not significantly different from zero. Thus, the choice of the comparison group seems to matter for the effect on the inactivity rate. Therefore, one should probably be very careful when making predictions about the effect of a third year of vocational studies on the future rate of inactivity.

6 Concluding remarks

In this paper I argue that conditional on a set of local characteristics the assignment of the pilot scheme, which preceded the 1991-reform of upper secondary education, is random. Thus, this gives me the opportunity to analyse the random variation in education that the pilot scheme gave rise to.

The paper begins with a reduced form analysis of the effects of living in a pilot municipality in 1990. The school leavers should have completed a two-year upper secondary vocational education in 1992 or a three-year upper

[20] Adding controls for county of residence and local unemployment rate at the time of the outcome the estimates for 1993 and 1995 are no longer statistically significant at the five per cent level.

[21] Estimating separate regressions for females and males the inactivity rate effect for the females turns out to be statistically insignificant for all years.

[22] It should be noted that an analysis of the effect on earnings has also been performed. The problem with this analysis is the short time period. However, the main result is that the estimates of the reduced form and instrumental variable equations are negative and imprecise for individuals who lived in a pilot municipality.

secondary vocational education in 1993. Living in a pilot municipality increases the probability of having a third year of upper secondary education by 4.5 percentage points. This effect is statistically significant and virtually constant throughout the time period. This is not surprising, since the three-year vocational programmes only were available in the pilot municipality. Still, this finding is important since pilot status has predictive value for a third year of vocational education at the upper secondary level.

One of the important reasons for adding a third year was to facilitate university enrolment for school leavers with a vocational education. The results show that the probability of university enrolment is positive for individuals living in a pilot municipality in 1990. A three-year upper secondary vocational education increases the probability of university enrolment by a third for individuals whose educational attainment was affected by living in a pilot municipality. The effect on university enrolment is large in magnitude. One possible explanation of this is that the increased focus on theoretical subjects in the three-year upper secondary vocational education gave increased possibilities of continuing to higher education.

The results on the inactivity rate suggest that the probability of being inactive is positive for individuals who had a three-year vocational education. The reform had two objectives: i) facilitate university enrolment; ii) reduce the exposure to unemployment. It thus seems that the reform was successful on the first, but not on the second account. One possible reason for the less favourable outcome with respect to the second policy objective is the increase in unemployment rate among young people between 1992 and 1993. The individuals with the two-year upper secondary vocational education leaving school in 1992 may have managed to establish themselves on the labour market before it declined. However, a sensitivity analysis was carried out using another comparison group, i.e. individuals who left the upper secondary school at the same time as individuals having a three-year vocational education. The effect of a third year in upper secondary school on the rate of inactivity changed from being significantly positive to being zero and statistically insignificant. Therefore, the effects on the inactivity, presented here, should be interpreted with due care. Nevertheless, it is fair to say that the reform did not reduce the vulnerability to unemployment.

References

Bennett, R et al. (1995) "Investing in skill: Expected returns to vocational studies", *Education Economics*, vol. 3, no. 2, pp. 99–117.

Card, D (1999) "The Causal Effect of Education on Earnings", in Ashenfelter O & D Card, (eds), *Handbook of Labour Economics*, pp. 1801–1863.

Cloonan, M & R Canning (2000) "Completion rates of Scottish vocational qualification (SVQ) courses: A research study", Scottish Educational Review, vol. 32, no. 1, pp. 55–67.

Denny, K & C Holm (2000) "Educational Policy Reform and the Return to Schooling from Instrumental Variables", CEPR Discussion paper 2518, London.

Edin, P-A, A Forslund & B Holmlund (2000) "The Swedish youth labor market in boom and depression Youth employment and joblessness in advanced countries" in Blanchflower D G & R B Freeman (eds), pp. 357–380.

Government Bill (1987/88:102) *Utveckling av yrkesutbildningen i gymnasieskolan*.

Government Bill (1990/91:85) *Växa med kunskaper – om gymnasieskolan och vuxenutbildningen*.

Gregg, P (2001) "The impact of youth unemployment on adult unemployment in the NCDS", *Economic Journal*, vol. 111, no. 475, pp. F626–F653.

Harmon, C & I Walker (1995) "Estimates of the Economic Return to Schooling for the United Kingdom", *American Economic Review*, vol. 85, no. 5, pp. 1278–1286.

Heraty, N, M J Morley & A McCarthy (2000) "Vocational Education and Training in the Republic of Ireland: institutional reform and policy developments since the 1960s", *Journal of Vocational Education and Training*, vol. 52, no. 2, pp. 177–98.

Imbens, G W & J D Angrist (1994) ”Identification and Estimation of Local Average Treatment Effects”, *Econometrica*, vol. 62., no. 2, pp. 467–475.

Johansson, E & L Strandberg (1999) Kan gymnasieskolan förse högskolan med studenter?, Omvärldsanalys för högskolan, Arbetsrapport nr 5, Högskoleverket, Stockholm. http://www.hsv.se/sv/CollectionServlet?page_id=655 (2002-11-14).

Lasonen, J (1996) *Reforming upper secondary education in Europe*, The Leonardo da Vinci Project Post–16 Strategies, Institute for Educational Research, Publication series B 92, University of Jyväskylä.

Lundahl, L (1997) *Efter svensk modell, LO, SAF och utbildningspolitiken 1944–1990*, Boréa Bokförlag, Umeå.

Meghir, C & M Palme (1999) “Assessing the Effect of Schooling on Earnings using a Social Experiment”, Working Paper W99/10, Institute for Fiscal Studies, London.

National Agency for Education (1999) ”Utvärdering av fem gymnasieprogram 1998”, rapport nr 163, Stockholm.

National Agency for Education (2000a) ”Utvärdering av fem gymnasieprogram 1999”, rapport nr 182, Stockholm.

National Agency for Education (2000b) ”Reformeringen av gymnasieskolan – en sammanfattande analys”, rapport nr 187, Stockholm.

National Agency for Education (2002) “The Swedish school system”, http://www.skolverket.se/fakta/faktablad/english/index.shtml (2002-08-22).

OECD (2002) *Education at a Glance*, Paris.

SOU 1989:90 *Utvärdering av försöksverksamhet med 3-årig yrkesinriktad utbildning i gymnasieskolan första året*, UGY, Rapport 1.

SOU 1989:106 *Sextusen platser och tiotusen platser för försök i gymnasieskolan Hur, var och varför?*.

SOU 1990:75 *Utvärdering av försöksverksamheten med 3-årig yrkesinriktad utbildning i gymnasieskolan*, UGY, Rapport 2.

SOU 1992:25 *Utvärdering av försöksverksamheten med 3-årig yrkesinriktad utbildning i gymnasieskolan*, UGY, Slutrapport.

SOU 1996:1 *Den nya gymnasieskolan – Hur går det?*, Utbildningsdepartementet (Ministry of Education).

SOU 1997:1 Den nya gymnasieskolan – Steg för steg, Utbildningsdepartementet (Ministry of Education).

SOU 1997:107 *Den nya gymnasieskolan – Problem och möjligheter*, Utbildningsdepartementet (Ministry of Education).

SOU 2002:120 *Åtta vägar till kunskap – En ny struktur för gymnasieskolan*, Slutbetänkande av Gymnasiekommittén 2000.

Statistics Sweden (1987) ”Gymnasieskolan 1986/87: Sökande och intagna höstterminen 1986”, Statistiska meddelanden U 50 SM 8701, Statistics Sweden, Örebro.

Statistics Sweden (1991) ”Gymnasieskolan 1990/91: Sökande och intagna höstterminen 1990”, Statistiska meddelanden U 50 SM 9101, Statistics Sweden, Örebro.

Statistics Sweden (2002) ”Sysselsättning och arbetslöshet 1975–2001”, Avdelningen för arbetsmarknads- och utbildningsstatistik.

Statistics Sweden, *Statistical Yearbook for Sweden,* various issues.

Wald, A (1940) “The fitting of straight lines if both variables are subject to error, *The Annals of Mathematical Statistics*, vol. 11, no. 3, pp. 284–300.

Appendix A – The dimension of the pilot municipalities

Table A1 The proportion of participating municipalities in the pilot scheme by county in the academic year 1990/91

County	Number of municipalities	Number of municipalities with vocational education	Number of pilot municipalities	Per cent of pilot muncipalities
Blekinge	5	5	4	80.0 %
Gotland	1	1	1	100.0 %
Gävleborg	10	7	7	100.0 %
Göteborgs and Bohus	15	9	8	88.9 %
Halland	6	5	2	40.0 %
Jämtland	8	7	5	71.4 %
Jönköping	11	7	5	71.4 %
Kalmar	12	6	3	50.0 %
Kopparberg	15	11	8	72.7 %
Kristianstad	13	10	5	50.0 %
Kronoberg	8	7	7	100.0 %
Malmöhus	20	10	7	70.0 %
Norrbotten	14	10	7	70.0 %
Skaraborg	17	10	9	90.0 %
Stockholm	25	21	9	42.9 %
Södermanland	7	5	4	80.0 %
Uppsala	6	4	4	100.0 %
Värmland	16	13	13	100.0 %
Västerbotten	15	7	6	85.7 %
Västernorrland	7	7	6	85.7 %
Västmanland	11	8	8	100.0 %
Älvsborg	18	10	8	80.0 %
Örebro	11	6	6	100.0 %
Östergötland	13	7	5	71.4 %
Total	**284**	**193**	**147**	**76.2 %**

Appendix B – Vocational programmes

Table B1 The vocational programmes before the reform and in the pilot scheme

The two-year vocational programmes before the reform	The three-year corresponding vocational programmes in the pilot scheme
Electrical and telecommunications engineering	Electrical and telecommunication engineering
Motor engineering	Motor and transport engineering
Metalwork	Industry
Process engineering	Productive engineering
Woodwork	Woodwork
Clothing manufacturing	Textile and clothing manufacturing
Building and construction	Building and construction
	Constructional metalwork
	Painting
	Heating and plumbing
Food manufacturing	Food manufacturing
	Restaurant
Nursing	Nursing
	Caring services
Community care	
Agriculture	Natural resource use
Forestry	
Gardening and landscaping	
Distribution and clerical	Commercial
Retail trade and clerical work	
Production and maintenance engineering	-
Consumer studies	-
-	Craft
-	Graphic

Note: The two-year community care programme also belonged to the three-year caring services programme.

Appendix C – Sample restrictions

Table C1 Number of individuals omitted due to sample restrictions

Restrictions	Pilot municipality		Non-pilot municipality	
	1992	1993	1992	1993
Total number of individuals	18 054	5 603	7 658	1 437
Number of individuals omitted due to				
non-comparable programme	1 834	66	586	8
two leaving certificates from upper secondary school	0	32	0	21
educational level equals compulsory education	105	0	1	0
not in employment register	0	0	0	1
grade point average equals zero	2	0	0	0
educational level not consistent with vocational programme	16	0	3	0
Total number of individuals in the analysis	**16 097**	**5 505**	**7 068**	**1 407**

Notes: Adding the year 1992 and 1993 within the pilot municipalities and within the non-pilot municipalities gives a total of 21 602 and 8 475 individuals, respectively. These numbers of individuals are not identical with the 1993 column of *Table 3* in Section 4.2, since the composition of the data changes as one move from 1992 to 1993.

Appendix D – Parameter estimates

Table D1 Reduced form estimates by outcome variable and year

	1993	1994	1995	1996	1997	1998
Years of upper secondary education (USE)						
Pilot municipality	0.0445	0.0484	0.0484	0.0469	0.0453	0.0455
	(0.0098)	(0.0099)	(0.0099)	(0.0099)	(0.0099)	(0.0100)
Female	-0.099	-0.085	-0.086	-0.086	-0.085	-0.086
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Immigrant	-0.034	-0.033	-0.030	-0.026	-0.030	-0.030
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Grade point average	0.079	0.074	0.077	0.078	0.078	0.079
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Unemployment rate 1990 in the municipality	0.749	0.619	0.758	0.725	0.795	0.742
	(0.566)	(0.570)	(0.571)	(0.571)	(0.572)	(0.574)
Proportion of highly-educated 1990 in the municipality	0.185	0.109	0.151	0.162	0.179	0.166
	(0.133)	(0.134)	(0.134)	(0.134)	(0.134)	(0.135)
Constant	0.223	0.281	0.272	0.268	0.269	0.266
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
University enrolment						
Pilot municipality	0.0049	0.0026	0.0039	0.0040	0.0088	0.0151
	(0.0035)	(0.0044)	(0.0051)	(0.0057)	(0.0062)	(0.0067)
Female	0.048	0.061	0.066	0.066	0.066	0.066
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Immigrant	0.011	0.021	0.027	0.023	0.023	0.030
	(0.005)	(0.006)	(0.007)	(0.008)	(0.009)	(0.009)
Grade point average	0.061	0.090	0.119	0.143	0.164	0.181
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Unemployment rate 1990 in the municipality	0.131	0.110	-0.160	-0.209	-0.316	-0.432
	(0.202)	(0.253)	(0.293)	(0.329)	(0.360)	(0.385)
Proportion of highly-educated 1990 in the municipality	0.013	0.097	0.203	0.282	0.331	0.375
	(0.047)	(0.059)	(0.069)	(0.077)	(0.084)	(0.090)
Constant	-0.178	-0.275	-0.366	-0.432	-0.485	-0.516
	(0.014)	(0.017)	(0.020)	(0.022)	(0.024)	(0.026)

Table D1 cont Reduced form estimates by outcome variable and year

	1993	1994	1995	1996	1997	1998
Inactivity rate						
Pilot municipality	0.0164	0.0084	0.0128	0.0077	0.0124	0.0113[23]
	(0.0076)	(0.0069)	(0.0056)	(0.0055)	(0.0055)	(0.0048)
Female	-0.097	-0.066	0.0007	0.018	0.030	0.030
	(0.005)	(0.004)	(0.0033)	(0.003)	(0.003)	(0.003)
Immigrant	0.088	0.099	0.085	0.090	0.084	0.061
	(0.011)	(0.0097)	(0.008)	(0.008)	(0.008)	(0.007)
Grade point average	-0.098	-0.079	-0.068	-0.070	-0.068	-0.046
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Unemployment rate 1990 in the municipality	0.977	0.976	1.107	0.164	0.254	-0.302
	(0.437)	(0.397)	(0.322)	(0.317)	(0.314)	(0.279)
Proportion of highly-educated 1990 in the municipality	-0.295	-0.303	-0.336	-0.349	-0.269	-0.141
	(0.102)	(0.093)	(0.075)	(0.074)	(0.074)	(0.065)
Constant	0.567	0.458	0.322	0.330	0.286	0.202
	(0.029)	(0.027)	(0.022)	(0.021)	(0.021)	(0.019)

Note: Standard errors are reported in parentheses. Regressions also include controls for county of residence and the composition of employment.

[23] The corresponding parameter estimate from the sensitivity analysis, i.e. using individuals with a two-year vocational programme completing the upper secondary education at the same time as the individuals with the three-year vocational programme is 0.0021, with a standard error of 0.0050.

Table D2 IV, OLS and probit estimates on university enrolment by year

	1993	1994	1995	1996	1997	1998
Instrumental variable estimates						
Three years of USE	0.111	0.055	0.081	0.084	0.194	0.332
	(0.083)	(0.092)	(0.107)	(0.123)	(0.143)	(0.161)
Female	0.059	0.066	0.073	0.073	0.082	0.094
	(0.009)	(0.008)	(0.010)	(0.011)	(0.013)	(0.014)
Immigrant	0.015	0.022	0.030	0.026	0.029	0.040
	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)
Grade point average	0.052	0.086	0.113	0.136	0.149	0.155
	(0.007)	(0.007)	(0.009)	(0.010)	(0.012)	(0.013)
Unemployment rate 1990 in	0.049	0.076	-0.221	-0.270	-0.470	-0.679
the municipality	(0.223)	(0.262)	(0.308)	(0.344)	(0.390)	(0.438)
Proportion of highly-educated	-0.0076	0.091	0.190	0.269	0.296	0.320
1990 in the municipality	(0.0522)	(0.061)	(0.071)	(0.080)	(0.091)	(0.102)
Constant	-0.202	-0.290	-0.388	-0.455	-0.537	-0.604
	(0.023)	(0.030)	(0.034)	(0.039)	(0.044)	(0.050)
OLS estimates						
Three years of USE	-0.014	-0.0053	0.0041	0.016	0.026	0.033
	(0.002)	(0.0026)	(0.0030)	(0.003)	(0.004)	(0.004)
Female	0.046	0.061	0.067	0.068	0.068	0.069
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Immigrant	0.011	0.020	0.027	0.024	0.024	0.031
	(0.005)	(0.006)	(0.007)	(0.008)	(0.009)	(0.009)
Grade point average	0.062	0.090	0.119	0.142	0.162	0.178
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Unemployment rate 1990 in	0.143	0.113	-0.162	-0.220	-0.335	-0.455
the municipality	(0.202)	(0.253)	(0.293)	(0.329)	(0.360)	(0.385)
Proportion of highly-educated	0.015	0.097	0.201	0.279	0.325	0.367
1990 in the municipality	(0.047)	(0.059)	(0.069)	(0.077)	(0.084)	(0.090)
Constant	-0.176	-0.274	-0.370	-0.438	-0.494	-0.528
	(0.014)	(0.017)	(0.020)	(0.022)	(0.024)	(0.026)

Table D2 cont IV, OLS and probit estimates on university enrolment by year

	1993	1994	1995	1996	1997	1998
Probit estimates						
Three years of USE	-0.0048	-0.0010	0.0061	0.017	0.026	0.034
	(0.0010)	(0.0016)	(0.0022)	(0.003)	(0.003)	(0.004)
Female	0.027	0.041	0.048	0.052	0.055	0.058
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
Immigrant	0.005	0.015	0.024	0.023	0.027	0.033
	(0.003)	(0.005)	(0.007)	(0.008)	(0.009)	(0.010)
Grade point average	0.030	0.061	0.094	0.126	0.154	0.177
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Unemployment rate 1990 in	0.072	0.0089	-0.258	-0.296	-0.388	-0.537
the municipality	(0.099)	(0.1714)	(0.227)	(0.284)	(0.331)	(0.370)
Proportion of highly-educated	0.014	0.070	0.172	0.258	0.312	0.352
1990 in the municipality	(0.022)	(0.038)	(0.050)	(0.064)	(0.074)	(0.083)

Notes: Standard errors are reported in parentheses. Regressions also include controls for county of residence and the composition of employment. The estimates from the probit model are the marginal effects. The marginal effect is the partial derivative of estimated probability with respect to the variable, evaluated at the means.

Table D3 IV, OLS and probit estimates on the rate of inactivity by year

	1993	1994	1995	1996	1997	1998
Instrumental variable estimates						
Three years of USE	0.369	0.174	0.265	0.164	0.275	0.249[24]
	(0.187)	(0.147)	(0.128)	(0.123)	(0.137)	(0.122)
Female	-0.061	-0.051	0.023	0.032	0.053	0.051
	(0.019)	(0.013)	(0.012)	(0.011)	(0.012)	(0.011)
Immigrant	0.101	0.105	0.093	0.094	0.093	0.069
	(0.013)	(0.011)	(0.010)	(0.009)	(0.010)	(0.009)
Grade point average	-0.127	-0.092	-0.088	-0.083	-0.090	-0.065
	(0.016)	(0.012)	(0.010)	(0.010)	(0.011)	(0.010)
Unemployment rate 1990 in the municipality	0.700	0.869	0.906	0.045	0.035	-0.487
	(0.500)	(0.418)	(0.371)	(0.345)	(0.374)	(0.332)
Proportion of highly-educated 1990 in the municipality	-0.363	-0.322	-0.376	-0.376	-0.318	-0.182
	(0.117)	(0.097)	(0.086)	(0.080)	(0.087)	(0.077)
Constant	0.485	0.409	0.250	0.286	0.211	0.135
	(0.051)	(0.048)	(0.041)	(0.039)	(0.043)	(0.038)
OLS estimates						
Three years of USE	0.020	0.0063	-0.0085	-0.017	-0.024	-0.023
	(0.004)	(0.0040)	(0.0033)	(0.003)	(0.003)	(0.003)
Female	-0.095	-0.066	0.00001	0.017	0.028	0.028
	(0.005)	(0.004)	(0.00333)	(0.003)	(0.003)	(0.003)
Immigrant	0.089	0.099	0.085	0.089	0.084	0.061
	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)	(0.007)
Grade point average	-0.100	-0.079	-0.067	-0.069	-0.066	-0.044
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Unemployment rate 1990 in the municipality	0.964	0.973	1.115	0.178	0.275	-0.283
	(0.436)	(0.397)	(0.322)	(0.317)	(0.314)	(0.279)
Proportion of highly-educated 1990 in the municipality	-0.301	-0.305	-0.337	-0.348	-0.267	-0.139
	(0.102)	(0.093)	(0.075)	(0.074)	(0.074)	(0.065)
Constant	0.559	0.454	0.321	0.333	0.289	0.205
	(0.029)	(0.027)	(0.0215)	(0.021)	(0.021)	(0.019)

[24] The corresponding parameter estimate from the sensitivity analysis, i.e. using individuals with a two-year vocational programme completing the upper secondary education at the same time as the individuals with the three-year vocational programme is 0.0399, with a standard error of 0.0950.

Table D3 cont IV, OLS and probit estimates on the rate of inactivity by year

	1993	1994	1995	1996	1997	1998
Probit estimates						
Three years of USE	0.020	0.0059	-0.0074	-0.016	-0.022	-0.021
	(0.004)	(0.0040)	(0.0031)	(0.003)	(0.003)	(0.003)
Female	-0.098	-0.067	-0.0014	0.015	0.026	0.027
	(0.004)	(0.004)	(0.0032)	(0.003)	(0.003)	(0.003)
Immigrant	0.092	0.100	0.080	0.084	0.078	0.057
	(0.013)	(0.012)	(0.010)	(0.010)	(0.010)	(0.009)
Grade point average	-0.098	-0.076	-0.061	-0.063	-0.060	-0.039
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Unemployment rate 1990 in	0.862	0.904	1.042	0.191	0.230	0.203
the municipality	(0.431)	(0.384)	(0.300)	(0.295)	(0.289)	(0.255)
Proportion of highly-educated	-0.301	-0.302	-0.329	-0. 332	-0.261	-0.130
1990 in the municipality	(0.101)	(0.092)	(0.075)	(0.072)	(0.072)	(0.064)

Notes: Standard errors are reported in parentheses. Regressions also include controls for county of residence and the composition of employment. The estimates from the probit model are the marginal effects. The marginal effect is the partial derivative of estimated probability with respect to the variable, evaluated at the means.

Essay IV: Earnings effects of adult secondary education in Sweden*

1 Introduction

Adult secondary education (ASE)[1] courses have been provided in Sweden since 1968. Despite this long tradition, we still know little about its labour market effects. ASE is part of the Swedish public sector school system and offers courses at the compulsory and upper secondary level, and also at a post-secondary training level. Courses at ASE aim to give adults who lack these types of education the equivalent qualification that young people obtain in the regular nine-year compulsory and upper secondary schooling system.

Until recently, there have not been any studies evaluating the labour market effects of ASE (Alm Stenflo, 2000; Westerlund, 2000; Axelsson & Westerlund, 2001, Stenberg, 2003). The growing interest of evaluating ASE is due to the large investment within ASE in the form of the "Adult Education Initiative" (AEI).[2] Westerlund (2000) and Axelsson & Westerlund (2001) estimate the effect on unemployment duration, whereas Alm Stenflo (2000) and Stenberg (2003) estimate the short-run earnings effects. Alm Stenflo (2000) finds positive earnings effects for individuals who completed ASE in 1992 or 1993 compared with non-participants. Stenberg (2003), on the other hand, finds negative annual earnings effects for individuals participating in AEI in the autumn of 1997, compared with those in labour market training.

The motives for participating in ASE may vary between individuals, but the motives probably include increasing the individuals' possibilities to change their situation in the labour market, e.g. through a change of occupation. Such

* I would like to thank Per-Anders Edin and Peter Fredriksson for valuable comments and discussions. Roope Uusitalo, and participants at seminars at IFAU and at Tinbergen Institute, Amsterdam have also provided useful comments. The usual disclaimer applies.

[1] The Swedish word is 'kommunal vuxenutbildning' and is often abbreviated 'komvux'. In this study I shall abbreviate it ASE.

[2] The AEI was introduced in July 1997, but came to an end in December 2002. The programme was aimed at unemployed people aged 25–55 who did not have three-year upper secondary education.

changes, in combination with a higher level of education, may in the long run also lead to changes in income levels.

This paper aims to provide the first evidence of the long-run earnings effects of participating in ASE. The long run is potentially important because, after participation in ASE some adults may enrol in higher education. The analysis is based on the Swedish Longitudinal Individual Data (LINDA) for the years 1983–2000. I examine the annual earnings effect in 2000 for those adults who enter a spell in ASE at some point between the spring of 1988 and the spring of 1993, compared with non-participants.

The remainder of the paper is organised as follows: Section 2 briefly describes the features of the ASE. The data and empirical approach is presented in Section 3. Section 4 first presents the model and then discusses the results. Concluding remarks are found in Section 5.

2 What is adult secondary education?

Adult education may be defined differently between countries. Therefore, this section presents, particularly for the non-Swedish reader, the institutional features of ASE in Sweden.

ASE is part of the Swedish public sector school system and is one form of adult education. Other forms are, for example, labour market training, staff training or qualified vocational education and training.[3] Apart from ASE, the public school system for adults includes education for adults with learning disabilities and Swedish for immigrants.

The basic idea behind ASE is to give adults lacking a three-year upper secondary education or its equivalent a second chance of obtaining this education. ASE, however, targets different groups of adults. In addition to the core group of adults with no compulsory or upper secondary education, the programme targets new entrants into the labour market. Adults who already are in the labour force and, for example, would like to take a course for further training in their profession/occupation, or adults who want to take the opportunity to change their occupation are two further targeted groups. A final group is people who take supplementary courses in order to qualify for university

[3] For further information about these, see National Agency for Education (2003) or Statistics Sweden (2001).

studies.[4] In reality, ASE is open to all individuals aged 20 and above. However, adults with the shortest education have priority over other adults for all courses. Applicants with a higher education than a two-year upper secondary education are admitted to the programme if there are vacant slots.

ASE offers courses at three educational levels: i) basic adult education, ii) upper secondary schooling and iii) post-secondary training. Basic adult education is equivalent to the comprehensive nine-year compulsory school and should give adults qualifications in the four core subjects: Swedish (or Swedish as a second language), English, mathematics and social studies. The adults passing these four core subjects will obtain a compulsory school leaving certificate. Subjects other than the four core subjects can also be included in the certificate. Adult upper secondary education is equivalent to the regular upper secondary education, but is not identical. The courses provided in the former may differ with regard to emphasis, content and scope. A certain level of completion in the core subjects must be reached in order to obtain a leaving certificate from the adult secondary education system. The post-secondary training programme provides the adults with further training in a specific occupation or training for a new occupation.

Individuals can apply to ASE twice a year: in the spring and in the autumn. Thus, individuals can begin a course in either term. The advantage of ASE is that participants can decide their own rate of study. Therefore, participants can combine their studies with employment, particularly since ASE is offered during the evenings as well as in the daytime.

Figure 1 shows the number of participating students[5] in ASE from 1970–2001. Enrolment is also presented separately for men and women.

[4] Another group has evolved during the last few years. This group consists of young adults (aged less than 25) taking ASE courses for improving their grades in order to qualify to university.

[5] A student is defined as a person who participates in ASE. In the statistics the person is counted as a student only one time, irrespective of how many courses the person is attending.

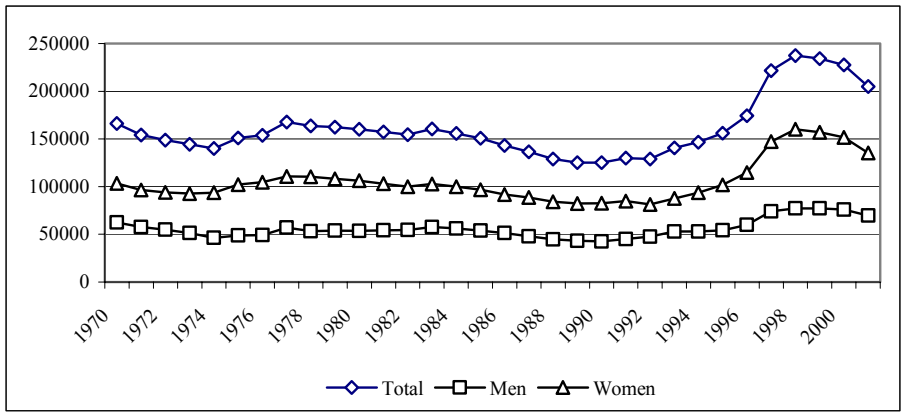


Figure 1 Students in ASE 1970–2001

Notes: The number of students is measured in the 42nd week in the years 1970–78. During 1979–91 the number of students is measured at the end of the term and, from 1992 and onwards the number is recorded in week 41 or 42 in October.

Source: Statistics Sweden, Statistical Yearbook for Sweden, various issues.

The dramatic change between 1996 and 1997 was the introduction of the Adult Education Initiative. Total enrolment was low between 1986 and 1993. During the whole period more women than men participated in ASE. In the United States too, more women seem to enrol in adult education than men (see for example Gilford, 1975; and Gill & Leigh, 2000).

3 Data and empirical approach

3.1 The data

The data used in this paper is the register-based Longitudinal Individual Data for Sweden (LINDA). LINDA consists of a representative sample of individuals and their household members. The individuals are the unit of observation in the present study. LINDA involves collection of different registers (see Edin & Fredriksson, 2000). In this study, I use the income register, local adult education register and higher education register. The crucial register is the local adult education register, which runs from 1988–95. This register contains information about which term the individual was registered for ASE, the level at which she studied and what course(s) read. Unfortunately, the register has no

information about whether the individual has completed or interrupted a course.[6] For this reason, participating in ASE should be referred to here as *registration* rather than completion.

To estimate the earnings effects of participating in ASE data on earnings are required some years before entering ASE. Since the local adult educational register begins in 1988, a pre-period of five years is included, which is the length of a normal business cycle. Therefore, the first year in my data is 1983. The last year of the period studied is 2000.

Before continuing to describe the participants in the ASE and the comparison group, two features of the data should be mentioned. First, there was no collection of data on registration in the autumn of 1992. Data was collected in the spring of 1993. Thus, individuals who registered in the autumn of 1992 and in the spring of 1993 were all registered in the spring of 1993. Second, the information about the individual's highest level of educational attainment is observed from 1990 and onwards. Therefore, in order to have a value for this variable prior to 1990, I have imputed the 1990 year value backwards. Although there are data prior to 1990, the pre-data are seriously incomplete, in particular for higher education. Individuals often do not take out their degree, which could explain the incomplete higher educational data. 1990 provides me with the most reliable educational data, since the information is obtained from the census.

3.2 Participants in ASE

As already mentioned, I have information on individuals participating in ASE sometime between the spring 1988 and the autumn 1995. There are, however, four main restrictions that the individuals must satisfy in order to be included in the analysis.

The first restriction rests upon an institutional change. From the academic year 1993/94 and onwards, additional public funds were invested in "unemployment courses"; these courses were introduced as a consequence of a redistribution of resources from labour market training conducted by the public employment office to the regular educational system (Government Bill 1992/93:150). To avoid the inclusion of adults participating in ASE for unemployment reasons, only those who decided to commence their studies before

[6] According to official statistics the course drop out rate is between 14–19 per cent for the years 1988/89–92/93 (Statistics Sweden, 1990; 1991; 1992; 1993; 1994).

the autumn of 1993 are included. They are, to a greater extent, assumed to consider ASE as a choice of studies rather than an alternative to unemployment or labour market training. Individuals who decided to enter the ASE before the autumn of 1993, but who continued to study in the autumn of 1993 and the terms thereafter are included in the analysis. The reason for including these individuals is that the decision to enter ASE was taken before the autumn of 1993.

To make the second restriction easier to understand, consider the flow diagram in *Figure 2*. The figure represents different flows in ASE for two individuals, denoted A and B. The spring and the autumn spells are denoted S and A, respectively.

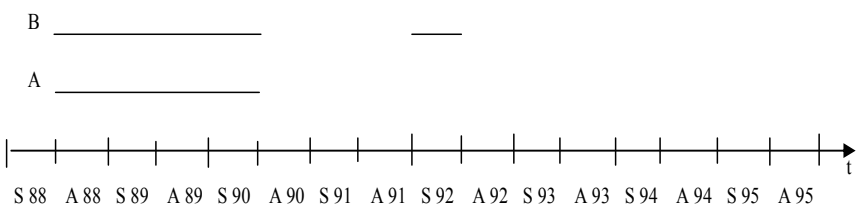


Figure 2 ASE inflows and outflows

Individual A and B both enter a first educational spell at the ASE in the autumn of 1988 and they participate in four terms terminating in the spring of 1990. In the spring of 1992 individual B embarks on a second spell. For both these individuals, the entry year in the ASE is 1988. Individual B's second educational spell is not included in the treatment group. I am only interested in the first educational spell, since this spell indicates the first decision to enter the programme. This means that each individual participating in ASE is only observed once, irrespective of the number of periods of ASE.

The third restriction is meant to alleviate the problems associated with the left-censoring of data in 1988. Preferably, one would prefer not to include individuals participating in a spell at ASE before 1988. Since data on participation in ASE are left-censored in 1988, I use auxiliary information. The information pertains to the use of study loans. I exclude individuals who have a posi-

tive study loan and who are not registered in higher education during the previous years 1983–87 for which I have data. This is not a perfect indicator of ASE participation, since it is possible to participate in ASE without taking a study loan. The assumption seems to be a good approximation for excluding potential participants in ASE before 1988, since, for example, 96 per cent of those who participated in the spring of 1992, had a positive study loan and were not registered in higher education during the previous years 1988–91.

Finally, the fourth restriction is that individuals must be between 25–55 years old when entering ASE. This restriction is because I am mainly interested in estimating the effect of ASE as a second chance for adults. The upper age limit of 55 years has been chosen due to the fact that individuals in Sweden may enter early retirement after this age. Therefore, participation in ASE after the age of 55 may not add anything extra to their working life.

I consider two different samples. Immigrants as a group may have a different motivation to commence ASE for example they need to learn Swedish to improve their possibilities to enter the Swedish labour market. As a result, immigrants who immigrated five years before entering ASE are excluded in the first sample. However, if the ASE is one of the key factors for immigrants to facilitate entry into the Swedish labour market, then this group should also be analysed. Thus, the second sample contains the immigrants who immigrated five years before commencing ASE.[7] The samples will be referred to the Swedish sample and the immigrant sample, respectively.

Appendix A presents the ASE characteristics of the participants for both samples. Male participants are found to be in upper secondary education, while female participants more frequently are at the post-secondary training level. It is evident that there is a change from participating more frequently in evening courses in the late 1980s towards participating more frequently in daytime courses during the 1990s. This change may be related to the Swedish recession. Language/linguistics is the most common main subject for male and female participants. Finally, the year 1993 differs from the other years, but this may only be due to the data collection mentioned earlier.

[7] In the LINDA database, an immigration year is given for all individuals who immigrate to Sweden. Thus, even a Swedish born-individual who immigrates to Sweden after a job abroad is allotted an immigration year. Since I am only interested in people who are born outside Sweden, additional information about the country of birth is used when sampling immigrants.

3.3 Design of the comparison group

In evaluation studies using non-experimental data, a generated comparison group is required, since one is not able to observe participants in the non-participant state. In this study, the comparison group consists of individuals who were never registered for ASE. It should be noted, however, that I can only be certain that the individuals included in the comparison group did not participate in ASE in 1988–95, for which data are available. Thus, this group of non-participants may consist of individuals who potentially have participated in ASE before 1988 and after 1995. However, to alleviate the problems associated with the inclusion of individuals who may have participated before 1988, the third restriction – information on study loans – is also adopted to the comparison group. Further, in order to lessen the extent that individuals who may have participated after 1995 are included in the comparison group the educational information in the year of 2000 is used.[8] One feature of educational information for this year is that it indicates whether the individual concerned has acquired his/her education as a result of ASE. Individuals having this educational attainment are dropped from the comparison group.[9]

Furthermore, the comparison group is, as is the case for participants, generated by selecting individuals aged 25–55 from each year during the period 1988–93, where each year corresponds to an entrance year. Thus, the same individual is selected every year, unless the individual fails to meet the age criteria. In other words, a non-participant occurs in the data a maximum of six times.[10]

[8] The year 2000 is the first year for the revised Swedish standard classification of education. The revision had two main purposes: i) to adjust the Swedish standard to the international standard for classification of education (ISCED 97), ii) to generate greater clearness and a more functional system than the old one (Statistics Sweden, 2000).

[9] However, the information suffers from problems, since one of the educational attainments – supplementary courses in upper secondary school – also covers those individuals who attend supplementary courses in the regular upper secondary school system. The individuals should not be eliminated, unless they received their qualification in the year 2000. This is the best possible way of eliminating potential participants after 1995, but since the educational information suffers from problems, I also included these individuals to see whether the results changed or not. Only small differences are found.

[10] The individual is repeated in the sample since he/she may commence in different entrance years. The observations from the same individual with different entrance year are correlated. To correct for this I use cluster-adjusted standard errors. For a documentation of the cluster-adjusted standard error, see Rogers (1993) and Williams (2000).

Since ASE is targeted at specific groups of adults, selection into the programme is present. For example, individuals with a two-year upper secondary education or less are selected into ASE to a greater extent, since this is one of the target groups. In this respect, participants differ from non-participants in terms of the level of education. However, participants may also differ in other respects, such as age.

In order to match non-participants to participants on the observed characteristics that seem to induce selection into ASE, a weighting scheme is employed. The comparison group is weighted separately for each entrance year, so that age and the level of education match that of the treatment group (participants). The weighting approach is implemented using a number of cells. The cells are divided into six age intervals: i) 25–29, ii) 30–34, iii) 35–39, iv) 40–44, v) 45–49 and vi) 50–55, and four educational levels: i) compulsory education or less, ii) two-year upper secondary education, iii) three-year upper secondary education and iv) university education. Each cell contains a group of individuals in a particular age-interval and level of education. This group of individuals (a cell) represents a value of the particular characteristics. For example, if six per cent of the participants with a two-year upper secondary education are aged 25–29, whereas the corresponding figure is three per cent in the comparison group, the comparison group is weighted by a factor of two (six divided by three) in comparison with the participants with these characteristics in order to obtain the same distribution as the participants.

As a result of fewer observations in the immigrant sample, the cells are reduced to include five age cohorts (*v* and *vi* are combined) and three educational levels (*ii* and *iii* are combined). In addition, the weighting scheme is not applied separately for each entrance year. By examining the means of the explanatory variables included in the regressions, I can discern how well the weighting procedure has matched the comparison group to that of the participants. *Table 1* presents sample means for the Swedish sample. *Table B1* in Appendix B gives the corresponding results for the immigrant sample.

Table 1 Means and standard deviation of ASE participants and the comparison group for the Swedish sample

Variables	Male participants	Male comparison	Female participants	Female comparison
	Mean	Mean	Mean	Mean
Age, τ-1	34.5 (7.9)	34.5 (8.0)	36.3 (7.6)	36.3 (7.8)
Age squared, τ-1	1254.5 (578.0)	1257.8 (588.2)	1374.8 (564.9)	1382.1 (585.5)
Compulsory, τ-1	0.156	0.158	0.190	0.177
2-year upper secondary school, τ-1	0.361	0.374	0.384	0.388
3-year upper secondary school, τ-1	0.161	0.155	0.106	0.109
University, τ-1	0.239	0.227	0.244	0.236
Pre-programme earnings in τ-1	97 151 (54 664)	122 737 (72 883)	70 050 (40 283)	83 612 (50 350)
Pre-programme earnings in τ-2	88 179 (47 275)	106 675 (62 835)	61 688 (35 820)	72 288 (43 629)
Pre-programme earnings in τ-3	76 372 (39 567)	89 524 (50 910)	52 647 (30 263)	60 953 (35924)
Pre-programme earnings in τ-4	63 831 (33 240)	74 107 (42 264)	44 328 (25 472)	51 115 (29864)
Proportion of zero earnings in τ-1	0.032	0.027	0.031	0.043
Proportion of zero earnings in τ-2	0.025	0.025	0.039	0.042
Proportion of zero earnings in τ-3	0.023	0.023	0.034	0.042
Proportion of zero earnings in τ-4	0.020	0.023	0.037	0.044
Proportion of entrance in ASE 1989	0.149	0.157	0.176	0.160
Proportion of entrance in ASE 1990	0.140	0.163	0.156	0.165
Proportion of entrance in ASE 1991	0.155	0.168	0.147	0.170
Proportion of entrance in ASE 1992	0.067	0.179	0.059	0.175
Proportion of entrance in ASE 1993	0.169	0.180	0.117	0.174
Unmarried, τ-1	0.516	0.530	0.329	0.352
Divorced, τ-1	0.070	0.057	0.112	0.096
Widow(er), τ-1	0.003	0.002	0.008	0.008
Nordic, τ-1	0.030	0.031	0.046	0.044
Other immigrants, τ-1	0.062	0.039	0.049	0.038
No. of observations	2 645	262 590	5 752	228 822

Notes: Standard deviation is in parentheses. τ is the ASE entrance year.

The weighting procedure seems to have worked rather well. The low proportion of about 0.06 in the entrance year 1992 is due to the collection of the data in 1992. The proportion of divorced people and the proportion of other immigrants are somewhat higher for the participants than for the comparison

group. This is the case both for males and females. The proportion of unmarried females is somewhat lower than in the comparison female group. Further, the table indicates that the proportion of zero earnings in the pre-programme year differs between participants and non-participants. Note that the participating females have a lower proportion of zero earners than non-participants. One would have expected the reverse pattern, that is, a higher proportion of zero earnings, as in the case of males.

Since selection by pre-programme earnings might be an issue, a second weighting scheme is constructed. The second weighting scheme therefore adds the information on the individual's pre-programme earnings, which are defined as an indicator of zero earnings. The number of cells increases and some cells are empty as a result, if this information is added. For the Swedish sample, I therefore aggregate age-intervals to only include two age-intervals: i) 25–34 and ii) 35–55. The two- and three-years of upper secondary education programmes are also combined. *Table B2* and *Table B3* in Appendix B present the results using the second weighting scheme.

Another interesting aspect would be to see whether the Ashenfelter's (Ashenfelter, 1978) dip exists. The average relative earnings around τ (the ASE entrance year) is calculated to assess this. The participants' average relative income is the ratio of their earnings each year divided by the non-participants' weighted average earnings each year. The result is shown in *Figure 3*.

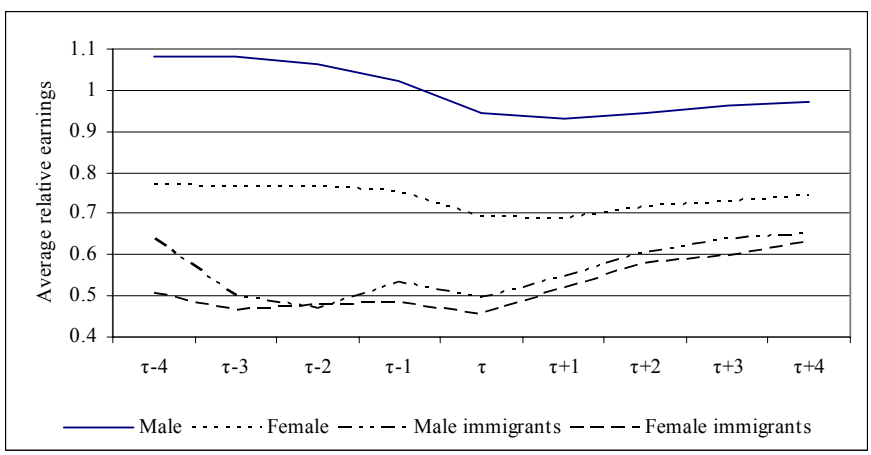


Figure 3 The participants' average relative earnings prior to, during and after entrance year in the ASE

The evidence is consistent with Ashenfelter's dip (Ashenfelter, 1978). There is a decline in the pre-programme earnings. The drop in the participants' relative earnings prior to the ASE entrance year (τ) is pronounced for Swedish males. Since few immigrants might be observed the further away one goes from τ, the result for the immigrant sample should be interpreted with caution. Nevertheless, it seems that the male immigrants do enjoy an increase prior to commencement of ASE.

4 Model and results

4.1 The empirical model

Suppose that some individuals begin an educational programme at time τ and some individuals do not. Suppose also that I am interested in the earnings effect of the programme. The participants may differ from the non-participants in, for example, motivation. Now, motivation presumably influences both the probability of participating in ASE (the causing variable) and in earnings (the outcome variable), but is rarely observed in the data. A good predictor, which takes into account various individual characteristics such as motivation is pre-programme earnings. Pre-programme earnings may pick-up the differences between participants and non-participants. I consider the following model:

$$y_{ijt} = \delta D_{ij} + \beta X_{ij,\tau-1} + \sum_{s=1}^{4} \gamma_s y_{ij,\tau-s} + \sum_{s=1}^{4} \omega_s Z_{ij,\tau-s} + \alpha_\tau + \varepsilon_{ijt} \qquad \text{j=p, c} \qquad (1)$$

where j indicates the participation status, p for participant and c for the comparison group (non-participants), and τ is the ASE entrance year. y_{ijt} is the post-programme annual earnings in the year of 2000; D_{ij}=1 if j=p; $X_{ij,\tau-1}$ is a vector of individual characteristics one year prior to the programme entry; $y_{ij,\tau-s}$ is a vector of pre-programme annual earnings expressed in 2000 prices; $Z_{ij,\tau-s}$ is a vector of dummies taking the value one if pre-programme annual earnings are zero; α_τ is the time-fixed effect for entering the ASE and ε_{ijt} is the error term. δ is the parameter of interest representing the programme effect on annual earnings for the individual participating in ASE sometime during the period from spring 1988 to spring 1993.

It is worth noting that the equation is estimated with annual earnings in levels, including zero earnings. Investigating the data, there seem to be more

zeros in 2000 than in the pre-programme earning years in the 1980s. The aggregate increase in unemployment, which took place during the 1990s, may have been a factor influencing future employment opportunities, since some of those participating in ASE completed their studies during the economic recession in 1993. Furthermore, individuals may also continue into higher education, which implies to a higher extent that they have no income. Another explanation could be that adults in an older age group already have entered early retirement.

4.2 Results

After experimenting with different sets of the available explanatory variables, I decided to include the following variables: age, age squared, educational level, marital status, country of birth, county dummies and dummies for each entrance year in ASE and also four years of pre-programme earnings and the related dummies indicating whether the individual had zero earnings in these years. These explanatory variables form the basic specification. Note that in the Swedish sample I control for four years of pre-programme earnings, whereas only one pre-programme earnings year is controlled when the sample of immigrants is used. In this study, the programme effect should be interpreted as the effect of having participated on any occasion during the spring of 1988 to the spring of 1993.[11]

The estimates from the basic specification for males and females of the programme effect are presented in *Table 2*. The parameter estimates for the other variables included in the regressions are available in Appendix C. Note that columns (1) and (2) in *Table 2* presents the results using the weighting scheme (i) matched for age and level of education for each entrance year in ASE, while columns (3) and (4) re-estimate the programme effect using the weighting scheme (ii). The second scheme uses the pre-programme earnings (indicating one if the pre-programme earnings are zero, otherwise zero) as an additional matching variable for those in scheme (i).

[11] The duration of training is not taken into consideration. The reason for not doing so is that I only observe the number of terms, since an ASE course may, for example, only last for seven weeks, which is not a full term. In addition, it may also be endogenous.

Table 2 OLS estimates of the effect of adult participation in ASE on absolute earnings in the year of 2000, expressed in percentage terms

	Weighting scheme (i)		Weighting scheme (ii)	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
The Swedish sample				
Programme effect	-0.0374***	0.0028	-0.0356***	0.0038
	(0.0107)	(0.0074)	(0.0108)	(0.0075)
No. of participants	2 645	5 752	2 645	5 752
No. of non-participants	262 590	228 822	261 406	228 430
The immigrant sample				
Programme effect	0.0606	0.0905*	0.0483	0.0902*
	(0.0559)	(0.0492)	(0.0623)	(0.0520)
No. of participants	393	385	393	385
No. of non-participants	3 777	3 000	3 642	2 964

Notes: Weighting scheme (i): Age and level of education by entrance year in ASE, where the entrance year is applied only to the Swedish sample. Weighting scheme (ii): Adds the dummy for pre-programme earning as a matching variable. The basic specification controls for age, age squared, educational level, marital status, country of birth, county dummies, dummies for the ASE entrance year, pre-programme earnings and related dummies for zero earnings in the pre-programme earning years. Cluster-adjusted standard errors are in parentheses. Level of significance: *** = 1 per cent level, * = 10 per cent level.

Beginning with the native-born males and weighting scheme (i), the table shows that the effect of participating in ASE on earnings is significantly negative. For males, participating in ASE is associated with an earnings reduction of about 3.7 per cent.[12] In a recent study, Stenberg (2003) also finds a negative programme effect on earnings for those males who participated in the Adult Education Initiative (AEI) in the autumn of 1997 or 1998 compared with participants in labour market programmes. Stenberg (2003) reports a short-run negative effect of about SEK 19 000 using the outcome year of 1999 and 2000, respectively.[13] Leigh & Gill (1997) estimates the returns to community college. Males and females are separated into returning adults and continuing students.

[12] Expressed in absolute terms this is about SEK 9 500.

[13] When Stenberg (2003) used the outcome year 2000 for those participating in the autumn of 1997, the negative effect was reduced to about SEK 12 000.

They find a more positive effect on earnings for the returning males than males continuing to community college directly from upper secondary school studies.

No programme effect is found for the native-born female participants. This result is in line with Leigh & Gill (1997), who also find small statistically insignificant results. Stenberg (2003), on the other hand, reports a significantly negative short-run programme effect for the females.

What happens with the result when the dummy for pre-programme earnings is added as a matching variable in the weighting procedure? Columns (3) and (4) of *Table 2* present the results. The native-born male effect is somewhat less negative, whereas there is no change in the estimate for the native-born females. In general, the two weighting scheme produces quite similar estimates.[14]

The programme effect for immigrants is more on the positive side, however. The estimate for female immigrants is significant at the ten per cent level. This result may suggest that immigrants (at least females), who are one of the target group in ASE, benefit more from participating than not participating in ASE. One feature of the result could be that knowledge of Swedish, which usually is a key requirement for employment, may give ASE participants greater chances of obtaining a more qualified job than non-participants. A more qualified occupation should eventually result in higher earnings in the long-run.

It should be noted, however, that the above results are robust to inclusion of different variables. In general, adding zero earnings in the pre-programme year as a matching variable does not seem to affect the programme effect to any considerable extent. However, it seems that the native-born individuals with zero earnings in the pre-programme year benefit more from participating in ASE than other individuals, whereas this is not the case for the immigrants.

Would the estimate change if one adds an interaction between the dummy of having zero earnings in the pre-programme year and the participating dummy as a control, instead of weighting the comparison group with the former variable? This exercise is carried out as an assessment to the result using

[14] The results from a difference-in-difference technique, (a straightforward panel-data method) where the differences-in-differences estimator is the difference in earning growth between participants and non-participants, show somewhat different results. Using the pre-earnings year of 1985 and the post-earnings year of 2000 the programme effect is negative for both males and females at a magnitude of 9.6 per cent and 1.7 per cent, respectively. The results are statistically significant at the one per cent and five per cent level for males and the females, respectively. For males, this result implies that participating in ASE is still associated with negative earnings. For the females, participation in ASE is also associated with negative earning if this estimator is used, even though the negative effect is small.

weighting scheme (ii). Notice that this changes the interpretation of the effect: The programme effect should be interpreted as the effect for those with positive earnings in τ-1. For the Swedish sample, the estimate turns out to be the same as the estimates using weighting scheme (i). The estimate for female immigrants, changes marginally compared with scheme (ii), but is now statistically insignificant. The estimate for male immigrants is negative (-0.0322 log points) and is not statistically significant. This exercise shows that pre-programme earnings do not seem to matter.

It is puzzling to find that well-meaning programmes have a negative impact on individual outcomes. My result is, however, not new in Sweden. Several previous studies evaluating active labour market programmes appear to indicate that they are harmful for re-employment (see Larsson, 2003 among others). However, my result seems to be in line with a previous study by Dickinson *et al.* (1986) who evaluate the participation in the Comprehensive Employment and Training Act (CETA) on earnings. They found a negative significant impact on earnings for males and a modest positive earnings effect for females, although not at a statistically significant level. However, in a later study by the same authors (Dickinson *et al*., 1987) they find a statistically significant positive effect even on females' earnings, which results from the fact that females who participated in CETA are, partly, more likely to be employed, but also because of a positive effect on hours worked per week and weeks worked per year among those who worked. Even though my results seem to be in line with previous research, one wonders what the potential explanation is for the negative effect for Swedish males. In this paper, individuals who also participate at the compulsory level are included. However, if individuals who participate at the compulsory level in ASE are eliminated and the sample is re-weighted with the new population distribution, the result still shows a negative earnings effect for males.

4.3 Sensitivity analysis of the basic results

In order to asses whether the findings are robust, this section provides further elaboration of the basic specification. The first sensitivity analysis concerns whether the introduction of interaction terms changes the basic results. Two sets of interaction terms are considered, but only one at a time. The first is a set of interaction terms between the fixed effects entry year (α_τ) with the dummy for pre-programme earnings ($DINC_{\tau-1}$). The second is a set of interactions between α_τ and pre-programme earnings ($INC_{\tau-1}$). These interaction terms are

introduced because the programme may have different effects in different stages of the business cycle. In addition, there may be different selection into the programme depending on the cycle. The interaction terms should take care of such concerns. The results of the basic specification, adding the interaction terms as controls, are shown in *Table 3*.

Table 3 OLS estimates of the effect of adult participation in ASE on absolute annual earnings in 2000, expressed in percentage terms

	Weighting scheme (i)		Weighting scheme (ii)	
	Male	Female	Male	Female
The Swedish sample				
BS adding interaction term $\alpha_\tau * DINC_{\tau-1}$				
Programme effect	-0.0370***	0.0014	-0.0349***	0.0040
	(0.0107)	(0.0073)	(0.0108)	(0.0075)
BS adding interaction term $\alpha_\tau * INC_{\tau-1}$				
Programme effect	-0.0315***	0.0027	-0.0292***	0.0046
	(0.0107)	(0.0074)	(0.0108)	(0.0075)
The immigrant sample				
BS adding interaction term $\alpha_\tau * DINC_{\tau-1}$				
Programme effect	0.0632	0.0926*	0.0507	0.0922*
	(0.0562)	(0.0495)	(0.0623)	(0.0522)
BS adding interaction term $\alpha_\tau * INC_{\tau-1}$				
Programme effect	0.0607	0.0854*	0.0531	0.0865*
	(0.0567)	(0.0493)	(0.0628)	(0.0522)

Notes: Weighting scheme (i): Age and level of education by entrance year in ASE, where the entrance year is applied only to the Swedish sample. Weighting scheme (ii): Adds the dummy for pre-programme earning as a weighting variable. The basic specification (BS) control for age, age squared, educational level, marital status, country of birth, county dummies, dummies for the ASE entrance year, pre-programme earnings ($INC_{\tau-1}$) and dummies for zero earnings in pre-programme earning years ($DINC_{\tau-1}$). Cluster-adjusted standard errors are in parentheses. Level of significance: *** = 1 per cent level, * = 10 per cent level.

Introducing interaction terms does not change the results much. Thus, the estimates of the basic specification seem to be robust. The general findings for the Swedish sample are that estimates for males are less negative, whereas no changes appear for females. In general, the estimates for both male and female immigrants are somewhat more positive compared with the basic specification.

Another interesting feature would be to see whether the programme effect differs by age. To investigate this issue, the sample is split into young adults (25–42 years of age) and older adults (43–55 years of age). In a recent study,

Hill (2001) finds that when women participated in post-school-age training at an age of 30–44, the wage effect of training was higher at a higher age.

Table 4 presents the results. This analysis is only performed on the Swedish sample, since most of the individuals in the immigrant sample contain young adults.

Table 4 OLS estimates of the effect of adult participation in ASE for the Swedish sample using two age-intervals on absolute annual earnings in 2000, expressed in percentage terms

	Weighting scheme (i)		Weighting scheme (ii)	
	Male	Female	Male	Female
Aged 25–42				
Programme effect	-0.0285**	0.0002	-0.0262**	-0.0006
	(0.0132)	(0.0097)	(0.0132)	(0.0095)
Aged 43–55				
Programme effect	-0.0575***	-0.0002	-0.0629***	0.0003
	(0.0107)	(0.0113)	(0.0191)	(0.0120)

Notes: Weighting scheme (i): Age and level of education by entrance year in ASE. Weighting scheme (ii): Adds the dummy for pre-programme earning as a matching variable. The basic specification controls for age, age squared, educational level, marital status, country of birth, county dummies, dummies for the ASE entrance year, pre-programme earnings and related dummies for zero earnings in the pre-programme earning years. Cluster-adjusted standard errors are in parentheses. Level of significance: *** = 1 per cent level, ** = 5 per cent level.

The results in *Table 4* suggest that adult males who train at a young age, i.e. 25–42 have a lower earnings reduction than adult males entering training at a higher age (43–55 years of age). This implies that if males should participate in ASE, they should do so when they are young. No evidence is found for the females, however.

Since it is likely that participation in different courses may affect ASE participants differently I also investigate the programme effect if the main subject as a control is added. The estimated programme effect of course depends on the reference group. Therefore, all programme effects using the various main subjects as reference groups are presented. The analysis is only performed using the Swedish sample. The results are displayed in *Table 5*. Note that the main subject "art", which also include music and drama, is dropped due to too few observations.

Table 5 OLS estimates of the effect of adult participation in ASE for the basic specification of the Swedish sample adding main subject as a control for absolute annual earnings in 2000, expressed in percentage terms

Programme effect when each of the following main subject is the reference group	Weighting scheme (i)		Weighting scheme (ii)	
	Male	Female	Male	Female
Behavioural science, art subjects	0.0740	0.0024	0.0686	-0.0020
	(0.0472)	(0.0306)	(0.0468)	(0.0308)
Business economics, administration	-0.0160	0.0256	-0.0098	0.0311
	(0.0345)	(0.0227)	(0.0348)	(0.0232)
Maths, science	-0.0394	0.0323	-0.0400	0.0333
	(0.0255)	(0.0257)	(0.0256)	(0.0262)
Medicine, health services, nursing	-0.2258***	0.0227	-0.2219***	0.0191
	(0.0831)	(0.0277)	(0.0839)	(0.0281)
Social science, information technology	-0.0676***	-0.0052	-0.0636***	-0.0014
	(0.0239)	(0.0161)	(0.0241)	(0.0164)
Language and linguistics	-0.0313*	0.0067	-0.0302	0.0069
	(0.0183)	(0.0119)	(0.0186)	(0.0121)
Technology	-0.0358	-0.0565	-0.0323	-0.0575
	(0.0219)	(0.0363)	(0.0220)	(0.0370)
No related subjects	-0.2075**	-0.0252	-0.2105**	-0.0245
	(0.0823)	(0.0158)	(0.0825)	(0.0161)

Notes: Weighting scheme (i): Age and level of education by entrance year in ASE. Weighting scheme (ii): Adds the dummy for pre-programme earning as a matching variable. The basic specification controls for age, age squared, educational level, marital status, country of birth, county dummies, dummies for the ASE entrance year, pre-programme earnings and related dummies for zero earnings in the pre-programme earning years. Cluster-adjusted standard errors are in parentheses. Level of significance: *** = 1 per cent level, ** = 5 per cent level, * =10 per cent level.

Following addition of the main subject as a control, the basic pattern still seems to apply. That is, the programme effect for males are still negative and in some cases even statistically significant; whereas no evidence is found for the females. Furthermore, the males' distribution of main subjects differs from the females' distribution of main subjects. Is it the male distribution over subjects that give the negative estimate for males? Would females have a similar negative effect had they chosen "male" subjects? The answer to the latter question is no. The weighted average of the female estimate, using the male distribution as weights, is still positive. Thus, differences in the gender distribution over main subjects do not seem to drive the different point estimates for males and females result.

5 Concluding remarks

Very little attention has been paid to the labour market effects of participation in adult secondary education (ASE), although the programme has been offered since 1968. The only two previous studies focus solely on the short-run earnings effects of participating in the ASE (Alm Stenflo, 2000; Stenberg, 2003).

The short run effects may be very misleading as a basis for evaluation, since one of the objectives of the program is to prepare for higher education. Consequently, the purpose of this paper has been to provide the first evidence of the long-run earnings effects of participating in ASE. The analysis is based on a unique database; the Longitudinal Individual Data for Sweden. The analysis estimates earnings effects for males and females and also for native-born persons and immigrants.

The main results are the following. In the case of native-born males, participation in ASE is associated with a reduction in earnings. Their reduction is around three and a half per cent. No evidence is found for the native-born females, however. The programme effect for the female immigrants is, on the other hand, positive and significant on the ten per cent level. For the female immigrants, the earning increase is around nine per cent. These findings appear to be robust to reasonable alternative specifications. For males, the size of the programme effect depends on age. The effects are less negative for young adults and more negative for old adults compared with the main result. For old male adults, the reduction is almost six per cent.

Thus, the earnings effects of ASE are rather dismal, at least for the native-born population. This may suggest that the value of larger-scale interventions such as the Adult Education Initiative (AEI) is limited. It seems to me that the long-run effects of the AEI are an important topic for future research.

References

Alm Stenflo, G (2000) ”Inkomst- och sysselsättningseffekter av kommunal vuxenutbildning”, Temarapport 2000:1, Prognosinstitutet, Statistics Sweden, Stockholm.

Ashenfelter, O (1978) “Estimating the Effect of Training Programs on Earnings”, *Review of Economics and Statistics*, vol. 60, no. 1, pp. 45–57.

Axelsson, R & O Westerlund (2001) ”Inflöde och varaktighet som arbetssökande efter avslutad utbildning – Deltagare i Kunskapslyftet och arbetsmarknadsutbildning höstterminen 1997”, Umeå Economic Studies no. 566, Umeå University, Umeå.

Dickinson, K P, T R Johnson & R W West (1986) “An analysis of the impact of CETA programs on participants’ earnings”, *Journal of Human Resources*, vol. 21, no. 1, pp. 64–91.

Dickinson, K P, T R Johnson & R W West (1987) “The impact of CETA programs on components of participants’ earnings”, *Industrial and Labor Relations Review*, vol. 40, no. 3, pp. 430–441.

Edin, P-A & P Fredriksson (2000) “LINDA – Longitudinal INdividual DAta for Sweden”, working paper 2000:19, Department of Economics, Uppsala University, Uppsala.

Gilford, D M (1975) “Statistical Snapshots of Adult Continuing Education”, *Journal of Higher Education*, vol. 46, no. 4, pp. 409–426.

Gill, A M & D E Leigh (2000) “Community college enrolment, college major, and the gender wage gap”, *Industrial and Labor Relations Review*, vol. 54, no. 1, pp. 163–181.

Government Bill (1992/93:150) *Kompletteringsproposition med förslag till slutlig reglering av statsbudgeten för budgetåret 1993/94*.

Hill, E T (2001) “Post-school-age training among women: training methods and labor market outcomes at older ages”, *Economics of Education Review*, vol. 20, no. 2, pp. 181–191.

Larsson, L (2003) “Evaluation of Swedish youth labour market programs”, *Journal of Human Resources*, vol. 38, no. 4, forthcoming.

Leigh D E & A M Gill (1997) “Labor Market Returns to Community Colleges: Evidence for Returning Adults”, *Journal of Human Resources*, vol. 32, no. 2, pp. 334–53.

National Agency for Education (2003) “The Swedish school system”, http://www.skolverket.se/pdf/faktablad/adult_edu.pdf (2003-01-17).

Rogers, W H (1993) “Regression standard errors in clustered samples”, *Stata Technical Bulletin Reprints*, vol. 3, pp. 88–94.

Statistics Sweden (1990) ”Kommunal vuxenutbildning 1988/89, Höstterminen 1988”, Statistiska meddelanden U 30 SM 9001, Örebro.

Statistics Sweden (1991) ”Kommunal vuxenutbildning 1988/90, Höstterminen 1989”, Statistiska meddelanden U 30 SM 9101, Örebro.

Statistics Sweden (1992) ”Kommunal vuxenutbildning 1990/91, Höstterminen 1990 och läsåret 1990/91”, Statistiska meddelanden U 30 SM 9201, Örebro.

Statistics Sweden (1993) ”Kommunal vuxenutbildning 1991/92, Höstterminen 1991 och läsåret 1991/92”, Statistiska meddelanden U 30 SM 9301, Örebro.

Statistics Sweden (1994) ”Kommunal vuxenutbildning 1992/93, Elever och kursdeltagare”, Statistiska meddelanden U 30 SM 9401, Örebro.

Statistics Sweden, *Statistical Yearbook for Sweden*, various issues.

Statistics Sweden (2000) “Swedish standard classification of education”, Reports on Statistical Co-ordination for the Official Statistics of Sweden, Örebro.

Statistics Sweden (2001) *Vuxenutbildningen i Sverige 2000*, Bulls Tryckeri-aktiebolag, Halmstad.

Stenberg, A (2003) *An evaluation of the adult education initiative relative labor market training*, dissertation thesis, Umeå Economic Studies no. 609, Umeå University.

Westerlund, O (2000) ”Inskrivningstid vid arbetsförmedlingen efter avslutade studier – Deltagare i Kunskapslyftet och arbetsmarknadsutbildning 1997”, Umeå Economic Studies no. 526, Umeå University, Umeå.

Williams, R L (2000) “A Note on Robust Variance Estimation for Cluster-Correlated Data”, *Biometrics*, vol. 56, pp. 645–646.

Appendix A – ASE characteristics of the participants

Table A1 ASE characteristics of the participants by entrance year, the Swedish sample

	Male						*Female*					
	1988	1989	1990	1991	1992	1993	1988	1989	1990	1991	1992	1993
Educational level in ASE												
Compulsory	15.1	20.3	22.2	22.9	26.0	34.5	19.6	20.3	20.0	22.6	23.4	31.5
Upper secondary	51.0	35.8	38.7	42.0	48.6	54.6	37.3	30.5	39.8	34.0	36.1	49.5
Post-secondary training	32.9	43.7	39.2	34.6	24.9	11.0	42.6	49.1	50.1	42.8	39.9	18.9
Unknown	0.9	0.3	0.0	0.5	0.6	0.0	0.5	0.1	0.1	0.6	0.6	0.2
Day or evening course in ASE												
Day	32.7	37.6	40.3	43.2	55.9	66.4	37.4	44.4	42.1	47.9	51.5	64.4
Evening	66.4	62.2	59.7	56.3	43.5	33.6	62.1	55.5	57.8	51.5	47.9	35.5
Unknown	0.9	0.3	0.0	0.5	0.6	0.0	0.5	0.1	0.1	0.6	0.6	0.2
Main subjects in ASE												
Behavioural science, art subjects	4.4	3.3	4.1	5.4	4.0	9.0	6.4	5.5	4.1	5.6	6.8	10.2
Art, music, drama	0.1	0.0	0.0	0.0	0.0	0.0	0.9	0.5	0.6	0.7	0.3	0.2
Business economics, administration	21.3	17.0	10.0	13.4	11.9	11.4	23.6	15.5	16.5	13.0	11.5	13.9
Maths, science	33.3	27.7	33.5	34.9	43.5	51.5	25.3	19.4	20.0	23.8	27.8	39.8
Medicine, health services, nursing	0.8	0.5	1.4	0.5	0.6	0.2	3.7	4.4	5.3	5.6	5.0	6.7
Social science, IT	28.3	33.3	34.3	31.5	34.5	41.2	30.3	27.9	29.8	28.4	32.5	37.3
Language and linguistics	54.3	40.9	45.7	52.0	53.1	70.5	50.7	42.6	40.8	48.9	49.1	60.7
Technology	21.7	16.2	14.1	12.4	6.8	9.8	5.4	3.1	3.3	4.6	6.8	7.0
No related subjects	2.7	1.5	0.8	1.2	1.7	1.3	16.0	17.8	14.7	12.4	9.5	6.1
Unknown							0.2	0.1	0.1	0.0	0.0	0.3
Mean number of main subjects	1.7	1.4	1.4	1.5	1.6	2.0	1.6	1.4	1.4	1.4	1.5	1.8

Table A2 ASE characteristics of the participants by entrance year, the immigrant sample

	Male						*Female*					
	1988	1989	1990	1991	1992	1993	1988	1989	1990	1991	1992	1993
Educational level in ASE												
Compulsory	54.9	67.7	75.5	62.3	58.8	84.9	54.2	62.5	46.7	67.2	70.8	80.3
Upper secondary	35.3	14.7	10.2	18.9	20.6	11.6	27.1	15.6	10.0	17.2	12.5	14.5
Post-secondary training	9.0	17.7	14.3	17.0	20.6	3.5	18.8	21.9	43.3	15.5	16.7	4.7
Unknown	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5
Day or evening course in ASE												
Day	43.1	44.1	40.8	49.1	58.8	79.1	47.9	56.3	66.7	48.3	58.3	83.9
Evening	56.9	55.9	59.2	49.1	41.2	20.9	52.1	43.8	33.3	51.7	41.7	15.5
Unknown	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5
Main subjects in ASE												
Behavioral science, art subjects	5.9	0.0	2.0	1.9	2.9	1.7	4.2	0.0	0.0	3.5	12.5	4.2
Art, music, drama	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Business economics, administration	15.7	2.9	0.0	1.9	5.9	2.9	12.5	6.3	6.7	3.5	0.0	2.1
Maths, science	45.1	20.6	26.5	18.9	35.3	32.6	20.8	18.8	16.7	15.5	16.7	33.2
Medicine, health services, nursing	0.0	2.9	0.0	0.0	2.9	0.6	2.1	0.0	0.0	3.5	0.0	3.1
Social science, IT	29.4	23.5	32.7	35.9	41.2	19.2	35.4	31.3	26.7	22.4	33.3	19.7
Language and linguistics	80.4	85.3	81.6	84.9	73.5	93.6	77.1	75.0	56.7	84.5	87.5	92.2
Technology	3.9	5.9	2.0	5.7	2.9	2.9	10.4	0.0	3.3	10.3	4.2	3.6
No related subjects	3.9	0.0	0.0	1.9	0.0	1.7	2.1	9.7	26.7	3.5	0.0	0.5
Mean number of main subjects	1.8	1.4	1.4	1.5	1.6	1.6	1.6	1.4	1.4	1.5	1.5	1.6

Appendix B – Descriptive statistics

Table B1 Means and standard deviations of the ASE participants and the comparison group for the immigrant sample using weighting scheme (i)

Variables	Male participants	Male comparison	Female participants	Female comparison
	Mean	Mean	Mean	Mean
Age, τ-1	31.2 (6.3)	31.3 (6.3)	32.5 (6.6)	32.6 (6.7)
Age squared, τ-1	1011.9 (444.3)	1017.9 (443.9)	1101.4 (466.0)	1107.7 (480.1)
3-year upper secondary school, τ-1	0.399	0.403	0.351	0.351
University, τ-1	0.321	0.318	0.312	0.310
Pre-programme earnings in τ-1	28 260 (41 024)	64 280 (69 422)	25 761 (35 846)	46 232 (50 260)
Proportion of zero earnings in τ-1	0.430	0.227	0.403	0.266
Proportion of entrance in ASE 1989	0.087	0.114	0.083	0.113
Proportion of entrance in ASE 1990	0.125	0.153	0.078	0.149
Proportion of entrance in ASE 1991	0.135	0.173	0.151	0.176
Proportion of entrance in ASE 1992	0.087	0.220	0.062	0.223
Proportion of entrance in ASE 1993	0.438	0.245	0.501	0.248
Unmarried, τ-1	0.445	0.417	0.210	0.216
Divorced, τ-1	0.061	0.057	0.073	0.075
Widow(er), τ-1	0.003	0.001	0.013	0.009
Other immigrants, τ-1	0.954	0.844	0.914	0.771
No. of observations	393	3 777	385	3 000

Notes: Standard deviation is in parentheses. τ is the ASE entrance year.

Table B2 Means and standard deviation of the participants in ASE and the comparison group for the Swedish sample using weighting scheme (ii)

Variables	Male participants	Male comparison	Female participants	Female comparison
	Mean	Mean	Mean	Mean
Age, τ-1	34.5 (7.9)	35.7 (8.5)	36.3 (7.6)	37.5 (8.7)
Age squared, τ-1	1254.5 (578.0)	1344.5 (644.3)	1374.8 (564.9)	1483.7 (663.3)
Compulsory, τ-1	0.156	0.143	0.190	0.158
2-year upper secondary school, τ-1	0.361	0.359	0.384	0.382
3-year upper secondary school, τ-1	0.161	0.169	0.106	0.113
University, τ-1	0.239	0.227	0.244	0.236
Pre-programme earnings in τ-1	97 151 (54 664)	124 482 (75 786)	70 050 (40 283)	84 704 (49 909)
Pre-programme earnings in τ-2	88 179 (47 275)	108 382 (65 289)	61 688 (35 820)	73 344 (43 286)
Pre-programme earnings in τ-3	76 372 (39 567)	91 234 (52 538)	52 647 (30 263)	61 984 (35 816)
Pre-programme earnings in τ-4	63 831 (33 240)	75 891 (43 573)	44 328 (25 472)	52 122 (29 828)
Proportion of zero earnings in τ-1	0.032	0.033	0.031	0.036
Proportion of zero earnings in τ-2	0.025	0.028	0.039	0.035
Proportion of zero earnings in τ-3	0.023	0.025	0.034	0.036
Proportion of zero earnings in τ-4	0.020	0.025	0.037	0.038
Proportion of entrance in ASE 1989	0.149	0.157	0.176	0.160
Proportion of entrance in ASE 1990	0.140	0.163	0.156	0.165
Proportion of entrance in ASE 1991	0.155	0.168	0.147	0.170
Proportion of entrance in ASE 1992	0.067	0.179	0.059	0.175
Proportion of entrance in ASE 1993	0.169	0.180	0.117	0.174
Unmarried, τ-1	0.516	0.498	0.329	0.329
Divorced, τ-1	0.070	0.063	0.112	0.102
Widow(er), τ-1	0.003	0.002	0.008	0.011
Nordic, τ-1	0.030	0.032	0.046	0.045
Other immigrants, τ-1	0.062	0.042	0.049	0.038
No. of observations	2 645	261 406	5 752	228 430

Notes: Standard deviation is in parentheses. τ is the ASE entrance year.

Table B3 Means and standard deviations of the immigrant sample using weighting scheme (ii)

Variables	Male participants	Male comparison	Female participants	Female comparison
	Mean	Mean	Mean	Mean
Age, τ-1	31.2 (6.3)	31.3 (6.3)	32.5 (6.6)	32.6 (6.6)
Age squared, τ-1	1011.9 (444.3)	1018.8 (448.9)	1101.4 (466.0)	1106.2 (475.2)
3-year USS, τ-1	0.229	0.172	0.190	0.177
University, τ-1	0.321	0.318	0.312	0.310
Pre-programme earnings in τ-1	28 260 (41 024)	47 738 (66 592)	25 761 (35 846)	37 659 (48 852)
Proportion of zero earnings in τ-1	0.430	0.430	0.403	0.403
Proportion of entrance in ASE 1989	0.087	0.105	0.083	0.114
Proportion of entrance in ASE 1990	0.125	0.144	0.078	0.148
Proportion of entrance in ASE 1991	0.135	0.162	0.151	0.170
Proportion of entrance in ASE 1992	0.087	0.227	0.062	0.225
Proportion of entrance in ASE 1993	0.438	0.276	0.501	0.255
Unmarried, τ-1	0.445	0.396	0.210	0.203
Divorced, τ-1	0.061	0.053	0.073	0.069
Widow(er), τ-1	0.003	0.001	0.013	0.008
Other immigrants, τ-1	0.954	0.875	0.914	0.800
No. of observations	393	3 642	385	2 964

Notes: Standard deviation is in parentheses. τ is the ASE entrance year.

Appendix C – Parameter estimates

Table C1 Parameter estimates for the Swedish sample from the basic specification using weighting scheme (i)

Variables	*Male*		*Female*	
	Parameter estimates	Cluster-adjusted standard error	Parameter estimates	Cluster-adjusted standard error
Programme effect	-9 528.705	2 729.085	475.5337	1 252.567
Age, τ-1	6 274.3	1 146.7	13 707.9	506.9
Age square, τ-1	-167.9	14.7	-216.2	6.3
Compulsory, τ-1	-30 250.9	2 814.3	7 519.7	1 590.0
2-year upper secondary school, τ-1	-17 902.9	3 226.0	15 284.4	1 387.7
3-year upper secondary school τ-1	26 490.2	4 044.5	38 068.3	2 230.8
University, τ-1	101 072.0	2 817.4	63 856.5	2 737.7
Pre-programme earnings in τ-1	1.029	0.147	0.936	0.138
Pre-programme earnings in τ-2	0.309	0.103	0.111	0.016
Pre-programme earnings in τ-3	0.055	0.075	0.030	0.069
Pre-programme earnings in τ-4	0.113	0.129	0.209	0.045
Dummy for zero earnings in τ-1	37 635.2	12 162.9	388.3	6 580.1
Dummy for zero earnings in τ-2	4 623.1	5 681.3	-441.8	1 576.1
Dummy for zero earnings in τ-3	-11 696.6	4 882.7	-3 546.0	1 701.9
Dummy for zero earnings in τ-4	-16 613.1	7 352.6	-1 935.1	1 789.9
Dummy for entrance in ASE 1989	-17 784.8	970.0	-8 278.6	549.9
Dummy for entrance in ASE 1990	-36 171.7	2 068.6	-20 814.6	1 281.1
Dummy for entrance in ASE 1991	-65 368.5	4 342.2	-39 872.8	2 852.8
Dummy for entrance in ASE 1992	-96 525.1	5 816.1	-56 750.4	3 592.0
Dummy for entrance in ASE 1993	-106 296.7	4 789.3	-64 756.2	2 610.7
Unmarried, τ-1	-18 524.4	2 399.6	-15 770.4	1 260.7
Divorced, τ-1	-20 213.2	2 675.6	-13 681.4	2 065.5
Widow(er), τ-1	-9 670.5	11 181.3	-8 419.9	3 937.7
Nordic, τ-1	-18 926.4	4 576.3	-12 451.3	2 130.8
Other immigrants, τ-1	-41 867.4	3 320.7	-33 153.7	2 581.3
Constant	157 264.0	23 004.7	-98 631.1	9 581.6
R^2-adjusted	0.1904		0.3302	
F-value	257.58		411.37	
Mean income	254 918.3		170 330.6	

Notes: The regression also controls for county dummies. τ is the ASE entrance year.

Table C2 Parameter estimates for the immigrant sample from the basic specification using weighting scheme (i)

Variables	*Male*		*Female*	
	Parameter estimates	Cluster-adjusted standard error	Parameter estimates	Cluster-adjusted standard error
Programme effect	9 649.262	8 902.603	10 947.29	5 956.828
Age, τ-1	-3 210.6	5 357.8	13 650.9	3 769.8
Age square, τ-1	-7.9	72.2	-208.1	53.5
3-year upper secondary school τ-1	18 476.9	7 818.4	30 322.2	6 538.5
University, τ-1	100 203.2	12 085.2	73 940.0	8 764.0
Pre-programme earnings in τ-1	0.899	0.120	0.664	0.093
Dummy for zero earnings in τ-1	7 099.7	9 987.7	-2 375.5	6 246.3
Dummy for entrance in ASE 1989	-14 480.2	7 444.9	-6 625.3	5 289.9
Dummy for entrance in ASE 1990	-30 045.0	11 300.9	-17 241.1	6 282.9
Dummy for entrance in ASE 1991	-53 234.0	14 123.9	-28 323.5	7 210.1
Dummy for entrance in ASE 1992	-69 469.5	14 848.8	-40 843.0	7 637.0
Dummy for entrance in ASE 1993	-66 075.6	14 508.3	-44 603.6	7 558.6
Unmarried, τ-1	-34 109.5	9 069.7	19 126.1	8 363.8
Divorced, τ-1	-10 961.4	14 164.8	9 879.8	13 111.8
Widow(er), τ-1	19 228.5	46 956.9	3 179.8	19 803.1
Other immigrants, τ-1	-66 239.2	17 840.7	-15 320.1	9 906.5
Constant	291 202.3	99 242.3	-116 579.2	63 816.7
R^2-adjusted	0.2448		0.2414	
F-value	7.20		9.30	
Mean income	159 145		121 028.8	

Notes: The regression also controls for county dummies. τ is the ASE entrance year.

Table C3 Parameter estimates for the Swedish sample from the basic specification using weighting scheme (ii)

Variables	*Male*		*Female*	
	Parameter estimates	Cluster-adjusted standard error	Parameter estimates	Cluster-adjusted standard error
Programme effect	-8 938.204	2 723.887	631.337	1 244.489
Age, τ-1	6 062.6	1 250.7	1 4826.6	503.9
Age square, τ-1	-164.4	15.7	-231.2	6.2
Compulsory, τ-1	-34 544.2	2 795.2	5 368.5	1 581.9
2-year upper secondary school, τ-1	-21 486.6	3 025.4	12 546.5	1 341.6
3-year upper secondary school τ-1	19 407.6	3 717.8	33 983.4	2 201.6
University, τ-1	92 179.5	3 075.9	59 242.0	2 517.6
Pre-programme earnings in τ-1	0.962	0.107	0.931	0.126
Pre-programme earnings in τ-2	0.374	0.118	0.093	0.035
Pre-programme earnings in τ-3	0.042	0.079	0.067	0.045
Pre-programme earnings in τ-4	0.111	0.121	0.190	0.042
Dummy for zero earnings in τ-1	45 098.1	11 245.9	8 024.6	6 166.0
Dummy for zero earnings in τ-2	10 152.1	7 013.8	1 391.5	1 530.3
Dummy for zero earnings in τ-3	-12 330.4	5 100.9	-2 342.9	1 667.7
Dummy for zero earnings in τ-4	-17 286.4	7 057.5	-1 513.3	1 801.3
Dummy for entrance in ASE 1989	-14 383.1	703.4	-7 792.8	440.4
Dummy for entrance in ASE 1990	-33 773.7	1 809.9	-18 956.7	1 080.9
Dummy for entrance in ASE 1991	-64 118.2	3 920.5	-37 409.5	2 439.7
Dummy for entrance in ASE 1992	-92 647.1	5 718.7	-53 374.5	2 995.9
Dummy for entrance in ASE 1993	-101 574.7	4 874.7	-60 096.7	2 146.9
Unmarried, τ-1	-19 060.8	2 288.5	-14 822.0	1 254.8
Divorced, τ-1	-19 607.7	2 550.6	-11 797.4	1 909.3
Widow(er), τ-1	-9 125.9	10 274.7	-4 400.1	3 701.8
Nordic, τ-1	-19 362.1	4 322.5	-12 756.3	2 086.3
Other immigrants, τ-1	-39 627.6	3 323.5	-32 369.4	2 545.5
Constant	164 210.9	24 957.7	-118 375.5	9 644.8
R^2-adjusted	0.2054		0.3333	
F-value	280.61		435.77	
Mean income	251 218.5		166 798	

Notes: The regression also controls for county dummies. τ is the ASE entrance year.

Table C4 Parameter estimates for the immigrant sample from the basic specification using weighting scheme (ii)

Variables	*Male*		*Female*	
	Parameter estimates	Cluster-adjusted standard error	Parameter estimates	Cluster-adjusted standard error
Programme effect	6 959.33	8 973.423	10 402.66	6 001.525
Age, τ-1	-1 952.6	4 910.6	12 260.9	3 537.9
Age square, τ-1	-21.5	66.2	-189.7	49.8
3-year upper secondary school τ-1	24 696.2	10 913.2	28 676.7	7 882.3
University, τ-1	87 198.8	11 583.3	65 999.8	8 529.6
Pre-programme earnings in τ-1	0.873	0.117	0.687	0.096
Dummy for zero earnings in τ-1	3 352.8	10 041.6	-3 584.3	6 499.3
Dummy for entrance in ASE 1989	-12 346.3	7 922.4	-10 143.7	5 620.6
Dummy for entrance in ASE 1990	-25 241.2	10 870.1	-23 076.1	6 672.4
Dummy for entrance in ASE 1991	-49 802.3	12 593.9	-31 465.1	7 548.4
Dummy for entrance in ASE 1992	-63 147.0	13 184.4	-44 705.9	7 810.5
Dummy for entrance in ASE 1993	-58 492.1	12 760.2	-48 548.9	7 821.4
Unmarried, τ-1	-34 283.5	8 958.6	18 390.3	8 130.1
Divorced, τ-1	-3 229.9	15 605.2	14 214.6	13 041.7
Widow(er), τ-1	11 484.8	53 986.5	-885.2	17 331.0
Other immigrants, τ-1	-60 594.3	17 084.0	-16 810.2	9 935.8
Constant	264 405.6	90 162.73	-80 239.4	60 596.4
R^2-adjusted	0.2400		0.2479	
F-value	7.37		9.54	
Mean income	144 013.8		115 358.7	

Notes: The regression also controls for county dummies. τ is the ASE entrance year.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapport

2003:1 Mörk Eva ”De arbetsmarknadspolitiska progammens effekt på den kommunala skolan”

2003:2 Runeson Caroline & Anders Bergeskog ”Arbetsmarknadspolitisk översikt 2000”

2003:3 Runeson Caroline & Anders Bergeskog ”Arbetsmarknadspolitisk översikt 2001”

2003:4 Calleman Catharina ”Invandrarna, skyddet för anställningen och diskrimineringslagstiftningen”

2003:5 Rooth Dan-Olof & Olof Åslund ”Spelar när och var någon roll? Arbetsmarknadslägets betydelse för invandrares inkomster”

2003:6 Forslund Anders & Bertil Holmlund ”Arbetslöshet och arbetsmarknadspolitik”

2003:7 Fröberg Daniela, Linus Lindqvist, Laura Larsson, Oskar Nordström Skans & Susanne Ackum Agell ”Friåret ur ett arbetsmarknadsperspektiv – delrapport 1”

2003:8 Olofsson Jonas ”Grundläggande yrkesutbildning och övergången skola arbetsliv – en jämförelse mellan olika utbildningsmodeller”

2003:9 Olli Segendorf Åsa ”Arbetsmarknadspolitiskt kalendarium II”

2003:10 Martinson Sara & Martin Lundin ”Vikten av arbetsgivarkontakter: en studie av den yrkesinriktade arbetsmarknadsutbildningen i ljuset av 70-procentsmålet”

2003:11 Hedström Peter, Ann-Sofie Kolm & Yvonne Åberg ”Social interaktion och arbetslöshet”

2003:12 Runeson Caroline ”Arbetsmarknadspolitisk översikt 2002”

2003:13 Ekström Erika ”Inkomsteffekter av kommunal vuxenutbildning”

Working Paper

2003:1 Fredriksson Peter & Per Johansson “Program evaluation and random program starts”

2003:2 Mörk Eva “The impact of active labor market programs on municipal services”

2003:3 Fredriksson Peter & Per Johansson "Employment, mobility, and active labor market programs"

2003:4 Heckman James & Salvador Navarro-Lozano "Using matching, instrumental variables and control functions to estimate economic choice models"

2003:5 Fredriksson Peter & Bertil Holmlund "Improving incentives in unemployment insurance: A review of recent research"

2003:6 Lindgren Urban & Olle Westerlund "Labour market programmes and geographical mobility: migration and commuting among programme participants and openly unemployed"

2003:7 Åslund Olof & Dan-Olof Rooth "Do when and where matter? Initial labor market conditions and immigrant earnings"

2003:8 Håkanson Christina, Satu Johanson & Erik Mellander "Employer-sponsored training in stabilisation and growth policy perspectives"

2003:9 Carneiro Pedro, Karsten Hansen & James Heckman "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice"

2003:10 Heckman James & Jeffrey Smith "The determinants of participation in a social program: Evidence from at prototypical job training program"

2003:11 Skedinger Per & Barbro Widerstedt "Recruitment to sheltered employment: Evidence from Samhall, a Swedish state-owned company"

2003:12 van den Berg Gerard J & Aico van Vuuren "The effect of search frictions on wages"

2003:13 Hansen Karsten, James Heckman & Kathleen Mullen "The effect of schooling and ability on achievement test scores"

2003:14 Nilsson Anna & Jonas Agell "Crime, unemployment and labor market programs in turbulent times"

2003:15 Hedström Peter, Ann-Sofie Kolm & Yvonne Åberg "Social interactions and unemployment"

2003:16 Ekström Erika "Earnings effects of adult secondary education in Sweden"

2003:17 Heckman James & Xuesong Li "Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000"

2003:18 Eckstein Zwi & Gerard J van den Berg "Empirical labor search models: A survey"

2003:19 Abbring Jaap H & Gerard J van den Berg "A simple procedure for the evaluation of treatment effects on duration variables"

Dissertation Series

2002:1 Larsson Laura "Evaluating social programs: active labor market policies and social insurance"

2002:2 Nordström Skans Oskar "Labour market effects of working time reductions and demographic changes"

2002:3 Sianesi Barbara "Essays on the evaluation of social programmes and educational qualifications"

2002:4 Eriksson Stefan "The persistence of unemployment: Does competition between employed and unemployed job applicants matter?"

2003:1 Andersson Fredrik "Causes and labor market consequences of producer heterogeneity"

2003:2 Ekström Erika "Essays on inequality and education"