Essays on the production of human capital

Erik Ø. Sørensen

Bergen, November 2005.

Dissertation submitted to the Department of Economics, Norwegian School of Economics and Business Administration, in partial fulfilment of the requirements for the degree of dr. oecon.

Acknowledgments

I am grateful to the Department of Economics at the Norwegian School of Economics and Business Administration for financial and moral support for this project, and to the Department of Economics at Queen's University for hosting me one year as a visiting research student. At both depertments I have had more stimulating and helpful discussions with too many people to mention by name.

My advisory comittee, Kjell G. Salvanes, Christopher Ferrall and Gerard van den Berg, has been extremely helpful and encouraging. Working with Kjell G. Salvanes and Oddbjørn Raaum on our common papers has also been an important learning experience.

Tor Jakob Klette was on my advisory comittee before he so untimely left us. This thesis is dedicated to his memory.

Introduction: Human Capital

That education is important for skills and human welfare is an idea that has grown into mainstream common knowledge. As recently as 1901, however, when B. Seebohm Rowntree published the first edition of his magisterial survey of the working classes in York, *Poverty: A Study of Town Life*, he did not mention education. At that time, however, education could not any more be ignored. An act of government had made access to elementary education free for all in 1891, and when Rowntree published a second edition the next year, a 'supplementary chapter' addressed some additional topics that needed attention: Among them were "public houses and Clubs", education and old age pensions (pp 362–445, Rowntree, 1902).

While undoubtedly the roots of modern human capital theory can be found in the writings of Adam Smith, St. Thomas Aquinas and Aristotle, those of us lacking scholarly training depend on more recent contributions. We know, however, that the basics of human capital theory was known to Anders Nicolai Kiær, director of Statistics Norway in the period 1877–1913, when he gave a lecture on "the economic value of a human life" to the Norwegian Association of Economists. Drawing on German authorities, Kiær noted that the value of a life could be calculated both from stream of costs associated with bringing children up to productive age and from the income above subsistence which flowed as a result of such investments.[1] He presented Table 1 which shows the an average life-cycle income profile for working class men and women, calculated the monetary costs of educational improvements, and noted how the increased income stream that would follow would have to be balanced against the monetary costs of education and foregone earnings. He suggested following a (German) dr. Engle in using a 4% discount rate – but shied away from the final calculation of the net present value of education since he regarded the statistical data at his disposal as too ridden with errors.

Kiær envisioned calculating the value of human capital from something very much like a national accounting perspective, assuming a rate of dis-

[1]Kiær briefly noted the possibility of using life insurance values, but did not pursue this strategy.

Alderen	Mænd			Kvinder		
	aarlig udgift.	aarlig indtægt	balance.	aarlig udgift.	aarlig indtægt	balance.
0-6	116	-	-116	116	-	-116
6-10	174	-	-174	174	-	-174
10-14	232	33	-199	232	33	-199
14-20	319	330	+11	290	330	+40
20-25	406	577,5	+171,5	319	396	+77
26-60	464	660	+196	348	396	+48
60-70	348	330	-18	290	264	-26
70 og derover	290	66	-224	261	49,5	-211,5

Table 1: **Cost- and income streams for men and women.** The first column is age, the second to fourth column are expenses, income and "net surplus" for men, the fifth to seventh columns repeat this information for women. Table 2 in Kiær (1891). Kiær made the calculations according to the method of a dr. Becker.

counting with little regard for how individuals themselves solved their choice problems. Alfred Marshall, however, has a an individualistic terminology and grasp on the problem that is very modern:

> ...the investment of capital in the rearing and early training of the workers of England is limited by the resources of parents in the various grades of society, by their power of forecasting the future, and by their willingness to sacrifice themselves for the sake of their children.
>
> ...The professional classes especially, while generally eager to save some capital *for* their children, are even more on the alert for opportunities of investing it *in* them. And whenever there occurs in the upper grades of industry a new opening for which an extra and special education is required, the future gains need not be very high relatively to the present outlay, in order to secure a keen competition for the post.
>
> But in the lower ranks of society the evil is great. For the slender means and education of the parents, and the comparative weakness of their power of distinctly realizing the future, prevent

them from investing capital in the education and training of their children with the same free and bold enterprise with which capital is applied to improving the machinery of any well-managed factory (p. 467, Marshall, 1920).

Marshall continues his discussion of the problem of underinvestment in human capital among the "lower ranks of society". He has earlier mentioned how the problem arises because the worker "remains his own property" and cannot be held as security, giving rise to financial constraints among the lower ranks and cumulative effects through the generations. When John R. Hicks wrote a new standard reference on economics in 1939, the notion of human capital was not something which needed special discussion. "Human capital" is noted in the index, and is introduced in passing in the text: "... increment or decrement in the value of prospects due to changes in people's own earning power (accumulation or decumulation of 'Human Capital'),..." (p. 178, Hicks, 1939). Hicks introduced the terms in a discussion about the problems of defining income under aggregate uncertainty, but he certainly does not seem to find "human capital" a new or particularly interesting idea.

Mincer (1958) and Becker (1964) are generally regarded as the ones responsible for the transfer of human capital theory into the post-WWII new and more mathematical economic science. They brought the formal tools of investment analysis to bear on individual educational choice, and from now on there is a recognisable and modern labour economics field with all the idealisations and stylised models that we have gotten used to. Recently, sequential models of schooling and career choices in environments with uncertainty have become feasible (Keane and Wolpin, 1997).

The question Marshall raised about financial constraints and intergenerational transmissions of inequality, however, remained strongly on the minds of empirical researchers. Lee Soltow, a professor of Economics at Ohio State University, came to Norway to examine the historical development of inequality. In 1965 he published *Toward Income Equality in Norway*, in which he proposed a methodology and a tradition of empirical inquiry that laid the groundwork for much of current research. Soltow introduced use of massive administrative databases in empirical labour economics. He examined 120

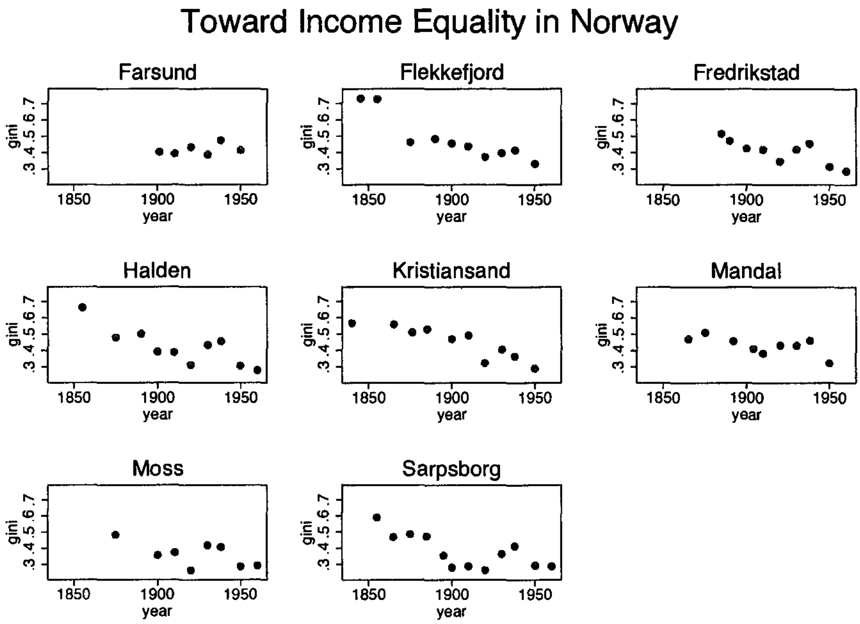


Figure 1: **Income inequality in eight Norwegian cities 1840–1960.** Data are taken from *exhibit 8* in Soltow (1965).

years of tax registers in 8 Norwegian cities. He sampled income data from these registers and calculated Gini-inequality indices for for the years 1840-1960. His main results, summarised in Figure 1, show a downward trend in urban inequality over this period.

Soltow discusses various structural changes to the Norwegian economy that can help explain this trend: changes in manufacturing, patterns of trade and socio-political changes. Of particular interest to labour economists is his discussion of education. He notes that greatly increased access to education probably helped reduce general inequality. But he also provides a specific analysis of the role of education in intergenerational processes. In Sarpsborg, one of the cities he studied, he was able to link people across time periods, and he got access to school records that included the name and address of parents. This very early use of linked administrative data helped him to establish an empirical relation between how well students do in school with their adult

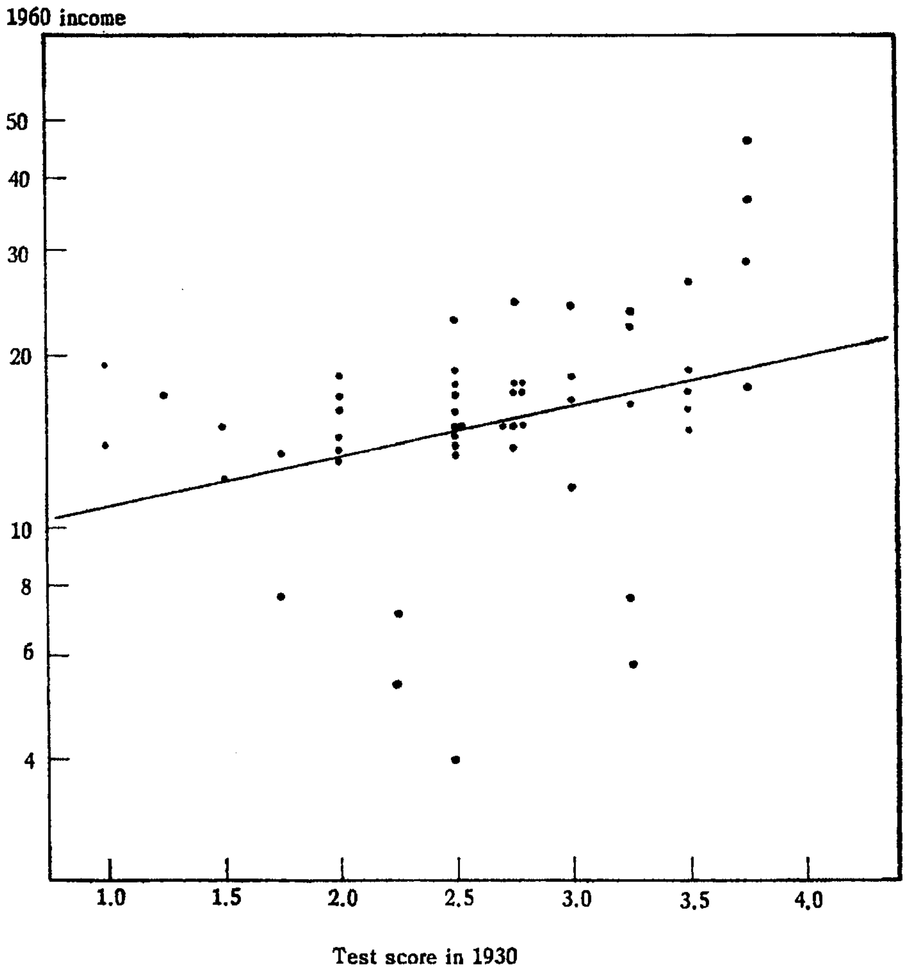


Figure 2: **Test scores and adult income in Sarpsborg.** This Figure is *exhibit 62* in Soltow (1965). Test score is from the end of 7th grade in 1930 while income is taken from the tax registers of the same people 1960.

	Son's median test score	
Father's income (y)	1930	1960
$y < P_{20}$	2.75	3.00
$P_{20} \leq y < P_{40}$	2.88	3.00
$P_{40} \leq y < P_{60}$	3.25	3.00
$P_{60} \leq y < P_{80}$	2.75	3.00
$P_{80} \leq y$	3.25	3.00

Table 2: **Fathers income and son's median test score.** Income measured in percentiles. Taken from Exhibit 61 in Soltow (1965).

earnings 30 years later, (Figure 2). This is not particularly surprising, but combined with information on how the relation between parental income and childrens educational outcomes has grown weaker (Table 2), he could provide a story in which changes in the workings of schools increase intergenerational mobility.

The work of Soltow, though today mostly forgotten, was an early attempt to make systematic use of linked administrative databases that were constructed for non-research purposes. This thesis adopts this strategy. I have been very fortunate to work with professor Kjell G. Salvanes on constructing a database of linked administrative records and census data at the Norwegian School of Economics and Business Administration. This work is made possible by an unique personal identifier given to each person, a process started in 1961-62 using the 1960 census (Skaug, 1968). We are therefore able to work with the full resident population instead of painstakingly constructing datasets from paper records such as Soltow was forced to do.

The three papers that make up this thesis take different approaches to the production of human capital. The first two, Raaum, Salvanes and Sørensen (2003, 2006), are in the tradition of Soltow, both in the use of data and in the reduced form approach. We study how family, local neighbourhoods and schools might have influenced adult earnings and educational outcomes. We find a small and declining role for the social environment (neighbourhoods and schools), but an important role for the family. These papers therefore indirectly support the methodological individualism adopted in the third paper on life-cycle career choices. This third paper adopts the full investment metaphor that Keane and Wolpin (1997) and related papers have taken from Becker and Mincer. My paper differs from Keane and Wolpin (1997) in modelling education as an input in an uncertain production of skills rather than taking education as a direct determinant of wages. This makes it possible to allow for sector-specific skills and for stochastic depreciation of skills that are not used.

The following sub-sections briefly characterise the papers of the thesis.

"The Neighbourhood is not what it used to be"

Co-authored with Oddbjørn Raaum and Kjell G. Salvanes. Forthcoming in *The Economic Journal.*

Using a variance decomposition framework that bounds the effect of families and neighbourhoods, we find important effects of family characteristics and residential location on adult education and earnings in Norway. Neighbourhoods are less important than families, as the correlations among siblings are significantly higher than among children growing up in the same local community. The impact of neighbourhoods is reduced by half from 1960 to 1970. We link this result to several policy changes in the 1960s aimed at increasing equality of opportunity in Norway. Neighbour correlations in Norway are found to be significantly lower than in the United States.

"The Impact of a Primary School Reform on Educational Stratification: A Norwegian Study of Neighbour and School Mate Correlations"

Co-authored with Oddbjørn Raaum and Kjell G. Salvanes. Published in *Swedish Economic Policy Review.*

School quality is hard to define and measure. It is influenced by not only school expenditures, but also characteristics that are hard to measure like norms and peer effects among teachers and pupils. Furthermore, family background and community characteristics are important in explaining educational outcomes. In this paper we study the composite effect of primary schools and neighbourhoods on adult educational attainment controlling for family characteristics. Instead of identifying the effect of specific neighbourhood and school characteristics on educational attainment, we focus on correlations in final years of schooling among neighbouring children and school mates. We find a clear trend of declining influence of childhood location over the 24 year period (birth cohorts 1947-1970). Then we ask whether a change in the compulsory school law extending the mandatory years of education, can explain this pattern. We find some effect of the primary school reform on the change in the neighbourhood effect. Motivated by the fact that neighbouring children typically go to the same school, we estimate school mate correlations for children born in the 1960s. The overall impact of factors shared by children who graduated from the same school at the age of 15/16 is negligible. The variation in "school quality" and the impact of peers on final educational attainment seem to have been very limited in Norway.

"Sectoral Choice with Human Capital and Accumulation of Pension Benefits"

Universal pension plans and large public-sector workforces affect accumulation and allocation of human capital. The benefit reforms and re-training programs being considered in many countries are likely to affect behaviour in ways that can only be analysed within forward-looking models of lifetime labour supply. Using Norwegian panel data on three birth cohorts, this paper develops and estimate a life-cycle model of public- and private-sector employment. The model of sequential career-choices builds on Keane and Wolpin (1997), extending the accumulation of skills to be sector-specific and allowing unobserved, non-deterministic depreciation of skills. The Norwegian retirement benefits are modelled in a way that builds on the discretisation approach of Rust and Phelan (1997), and identification is aided by exploiting how current career-choices affect future expected benefits. I find important heterogeneity in skill accumulation. The model is used to analyse the effect of a pension reform on sector-specific labour supply. The reform has large effects on labour supply, but the sectoral effects are small.

Errata to paper 1

On page "282" of paper 1, the condition in line 4-5 should be that

$$\operatorname{cov}(\widehat{\boldsymbol{a}}'\tilde{\boldsymbol{X}}_{cf}, \widehat{\boldsymbol{a}}'\tilde{\boldsymbol{X}}_{cf'}) \leq \operatorname{cov}(\boldsymbol{\alpha}'\tilde{\boldsymbol{X}}_{cf}, \boldsymbol{\alpha}'\tilde{\boldsymbol{X}}_{cf'}).$$

References

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education.* National Bureau of Economic Research, New York.

Hicks, J. R. (1939). *Value and Capital: An Inquiry Into Some Fundamental Principles of Economic Theory.* Clarendon Press.

Keane, M. P. and K. I. Wolpin (1997). "The career decisions of young men." *Journal of Political Economy,* **105**(3), 473–522.

Kiær, A. N. (1891). "Om menneskets økonomiske værd [on the economic value of a human life]." Lecture given to "Statsøkonomisk forening", 28th of December 1891. Printed in *Arr: Idéhistorisk tidsskrift,* 10(1):42-55, 1998.

Marshall, A. (1920). *Principles of Economics.* Macmillan Press, 8th edn. 1948 printing.

Mincer, J. (1958). "Investment in human capital and personal income distribution." *Journal of Political Economy,* **66**(4), 281–302.

Raaum, O., K. G. Salvanes and E. Ø. Sørensen (2003). "The impact of a primary school reform on educational stratification: A Norwegian study of neighbour and school mate correlations." *Swedish Economic Policy Review,* **10**(2), 143–170.

Raaum, O., K. G. Salvanes and E. Ø. Sørensen (2006). "The neighbourhood is not what it used to be." Forthcoming, *The Economic Journal*, January 2006.

Rowntree, B. S. (1902). *Poverty: A Study of Town Life.* Thomas Nelson & Sons, London, 2nd edn.

Rust, J. and C. Phelan (1997). "How social security and medicare affect retirement behavior in a world of incomplete markets." *Econometrica*, **65**(4), 781–831.

Skaug, H. (1968). "Oversikt over det sentrale personregister i statistisk sentralbyrå." i *Statistisk sentralbyrås sentrale registre*, Artikler fra Statistisk sentralbyrå 22.

Soltow, L. (1965). *Toward Income Equality in Norway.* University of Wisconsin Press.

The Economic Journal, 116 (*January*), 278–300. © Royal Economic Society 2006. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

THE NEIGHBOURHOOD IS NOT WHAT IT USED TO BE*

Oddbjørn Raaum, Kjell G. Salvanes and Erik Ø. Sørensen

Using a variance decomposition framework that bounds the effect of families and neighbourhoods, we find important effects of family characteristics and residential location on adult education and earnings in Norway. Neighbourhoods are less important than families, as the correlations among siblings are significantly higher than among children growing up in the same local community. The impact of neighbourhoods is reduced by half from 1960 to 1970. We link this result to several policy changes in the 1960s aimed at increasing equality of opportunity in Norway. Neighbour correlations in Norway are found to be significantly lower than in the US.

The role of families in determining socio-economic outcomes has been extensively discussed in economics as well as from other scientific perspectives. Parents influence their children via several channels: investment in their children's education, transmission of cultural values and social skills, and genetic endowments. Most studies show a strong degree of intergenerational transmission, since parental characteristics such as income and education are highly correlated with the outcomes of children along the same dimensions (Solon, 1999). Geographical location of the family is one specific aspect of parental behaviour. The close link between family and neighbourhood means that the impact of the two ought to be studied together. Since families self-select into neighbourhoods, resemblance in adult outcomes among childhood neighbours may reflect family rather than neighbourhood effects. But sibling resemblance can also be due to neighbourhood effects, since most brothers and sisters grow up in the same community. An unique Norwegian dataset provides the opportunity of an integrated and historical approach. We quantify the relative effects of families and neighbourhoods and examine their stability over time.

Neighbourhood effects is a label for a variety of different mechanisms. Some studies have focused on social interaction in peer-groups, through attitudes and preference formation as well as the existence and enforcement of social norms (Durlauf, 2001). Of course, neighbourhoods can also be important because of varying local resource bases, through availability of institutions such as schools and childcare. The significance of childhood location depends on whether these characteristics have any real impact and how they vary across neighbourhoods. Peer effects are likely to be amplified via sorting of advantaged families into advantaged neighbourhoods. The variation in local institutions such as schools

* We are grateful for comments from seminar participants at the Norwegian School of Economics and Business Administration, the University of Bergen, the University of Oslo, the University of Uppsala, Queen's University as well as helpful referees. Financial support was provided by the Norwegian Research Council, grant 120652/520 under the programme 'Competence, Education and Value Creation', under the 'Programme on Welfare Research', grants 140127/330 and 137236/530, and the "Programme on Efficiency in the Public Sector", grant 125251/520.

E	C	O	J			3	5	8	B	Dispatch: 7.11.05	Journal: ECOJ	CE: Blackwell
Journal Name				Manuscript No.						Author Received:	No. of pages: 23	PE: Revathi

and nurseries may also increase when sorting intensifies, and this would tend to reinforce peer effects.

The literature does not provide much evidence of neighbourhood effects on adult outcomes, and we know little about whether these effects are stable over time. The willingness of parents to pay a premium on house prices in order to get access to better schools and neighbourhood for their children shows that residential location is perceived as important (Black, 1999; Gibbons and Machin, 2003). Jargowsky (1996) finds some evidence of increased segregation by income in US metropolitan areas between 1970 and 1990.[1] However, the evidence supporting causal effects of childhood neighbourhood characteristics on opportunities or behaviour affecting adult outcomes is far from compelling, and identification is difficult. For instance, if we think of peer group influence as one potentially important channel of residential neighbourhood effects, it is hard to distinguish the selection into neighbourhoods from the impact of peer group behaviour on individual outcomes via social interaction. Most studies report unstable and small effects of community characteristics when these are included directly into the estimation equations of adult earnings or educational attainment (Ginther *et al.*, 2000). In addition to high correlation between family and neighbourhood characteristics due to sorting, it is also difficult to determine which characteristics to include and what biases result from using an incomplete set of characteristics.

In light of the difficulties of estimating causal effects of neighbourhood characteristics and previous failed attempts to disentangle them from family effects, we focus on the resemblance in adult earnings and educational attainment among siblings and neighbouring children (Solon *et al.*, 2000; Page and Solon, 2003).[2] If aspects of the family and/or residential community during childhood and adolescence are important determinants of adult outcomes, there will be a strong correlation between siblings as compared to two arbitrarily chosen individuals. It is possible to use this correlation to bound the share of neighbourhood effects in the total variance of outcomes.

In the present paper we focus on the effects of family background and childhood location on adult educational attainment and earnings. Rich data on the full population of Norwegian citizens enable us to construct neighbourhoods and families at the time of the 1960 and 1970 censuses, and to measure adult outcomes in 1990–5. We address the following questions:

What proportions of the variation in adult socio-economic outcomes, such as education and earnings, can be explained by family and childhood neighbourhood characteristics?

Are families more important than neighbourhoods?

Are these relationships stable over time?

[1] Kremer (1997) is sceptical of suggestions that changes in residential sorting will have a large impact on inequality in education and earnings. However, Kremer uses US census tracts as neighbourhoods, and these tracts are quite large. This may be one reason why he finds small effects of neighbourhood sorting.

[2] See Edin *et al.* (2003), Katz *et al.* (2001) and Oreopoulos (2003) for interesting attempts to use social experiment designs as an alternative to using data representative for a complete economy.

Our combined study of siblings and neighbouring children enables us to tell the extent to which Norwegian siblings' correlations reflect childhood location rather than common endowments or family environment. By comparing our estimates with those found in US data, we can also indicate whether these relationships are similar in different institutional settings.[3] Whether families and neighbourhoods have become more or less important determinants of adult outcomes is particularly interesting in light of the economic policies in Norway during the 1960s that aimed at increasing the equality of opportunity. Policies of the 1960s that may have influenced the effect of neighbourhoods include school reforms, increased access to student grants, and a radical redistribution of resources across municipalities.

The article is organised as follows. In the next Section we describe the statistical approach used to measure the effects of family and neighbourhood on adult education attainment and earnings. Section 2 contains a description of data as well as estimation procedures. The results are presented and discussed in Section 3, and Section 4 contains concluding remarks.

1. Statistical Model

Let y_{cfs} be an outcome variable, such as permanent income or years of education, for sibling s in the fth family in neighbourhood c. We assume that we can decompose additively the role of neighbourhood, family and individual-specific factors on the outcome as

$$y_{cfs} = \beta'\mathbf{Z}_c + \alpha'\mathbf{X}_{cf} + \epsilon_{cfs}, \tag{1}$$

where $\mathbf{X}_{cf}$ is a vector of all family characteristics that influence permanent earnings or years of education, $\mathbf{Z}_c$ contains all the neighbourhood characteristics, and ϵ_{cfs} represents individual factors orthogonal to both family and neighbourhood effects. Since $\mathbf{Z}_c$ and $\mathbf{X}_{cf}$ are latent vectors that include all relevant variables, it is not a restriction to let ϵ_{cfs} be orthogonal to the explanatory variables. We expect the family background effects, $\alpha'\mathbf{X}_{cf}$, and the neighbourhood effects, $\beta'\mathbf{Z}_c$, to be positively correlated. Advantaged families tend to sort into advantaged neighbourhoods and children in less favourable local communities tend to have parents with fewer resources.

We are looking for the relative influence of family and neighbourhoods on adult outcomes, $\mathrm{var}(\alpha'\mathbf{X}_{cf})/\mathrm{var}(y_{cfs})$ and $\mathrm{var}(\beta'\mathbf{Z}_c)/\mathrm{var}(y_{cfs})$. The relative variance of the neighbourhood effects, $\mathrm{var}(\beta'\mathbf{Z}_c)/\mathrm{var}(y_{cfs})$, measures the proportion of the variation in years of education or long-term earnings that can be attributed to factors shared by children from the same neighbourhood. Obviously, the correlation in adult outcomes between children growing up in the same local community cannot tell *why* neighbourhoods matter. They include the joint effects of variation in

[3] Studies from the US seem to dominate the existing evidence of family characteristics on adult outcomes, and especially the effect of community background has almost exclusively been studied within the US context. However, Ermisch and Francesconi (2001) and Dearden *et al.* (1997) use data for the United Kingdom to study the effect of family background. Gibbons (2002), Gibbons (2003) finds support for neighbourhood effects on educational attainment above family effects using UK data for 1970.

neighbourhood characteristics (Zs) and the causal impact they have on adult outcomes (β's).

If all the family and neighbourhood factors were observable and could be measured with accuracy, the strategy would be to estimate a regression model based on (1). Obviously, not all variables are observable or possible to measure with accuracy. Rather than arbitrarily choosing family and neighbourhood factors to include among those available, it is possible to bound the variance of neighbourhood effects by measuring neighbour covariances in y_{cfs} and observed family characteristics (Solon *et al.*, 2000). The total variance of the socioeconomic outcome can be decomposed as

$$\mathrm{var}(y_{cfs}) = \mathrm{var}(\boldsymbol{\beta}'\mathbf{Z}_c) + \mathrm{var}(\boldsymbol{\alpha}'\mathbf{X}_{cf}) + 2\mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \boldsymbol{\beta}'\mathbf{Z}_c) + \mathrm{var}(\epsilon_{cfs}). \tag{2}$$

The covariance between neighbouring children s and s' from families f and f is

$$\mathrm{cov}(y_{cfs}, y_{cf's'}) = \mathrm{var}(\boldsymbol{\beta}'\mathbf{Z}_c) + \mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \boldsymbol{\alpha}'\mathbf{X}_{cf'}) + 2\mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \boldsymbol{\beta}'\mathbf{Z}_c). \tag{3}$$

In (3), we notice that the neighbour covariance consists of more than the variance in neighbourhood characteristics given in the first term. The second term, $\mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \boldsymbol{\alpha}'\mathbf{X}_{cf'})$ represents sorting of similar families into the same areas, since neighbouring children come from similar families. We will assume that this covariance is positive. The third term, $\mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \beta'\mathbf{Z}_c)$, also represents sorting, in that it denotes the tendency of advantaged families to sort into advantaged neighbourhoods. We will assume that this is positive as well. We see that these two assumptions of positive sorting, together with the linear additive form of (1), imply that $\mathrm{var}(\beta'\mathbf{Z}_c) \leq \mathrm{cov}(y_{cfs}, y_{cf's'})$. The empirical neighbour covariance can therefore be interpreted as an upper bound on the variance of neighbourhood effects.

Children from the same family share both the neighbourhood and the family background,

$$\mathrm{cov}(y_{cfs}, y_{cfs'}) = \mathrm{var}(\boldsymbol{\beta}'\mathbf{Z}_c) + \mathrm{var}(\boldsymbol{\alpha}'\mathbf{X}_{cf}) + 2\mathrm{cov}(\boldsymbol{\alpha}'\mathbf{X}_{cf}, \boldsymbol{\beta}'\mathbf{Z}_c). \tag{4}$$

With the assumptions as outlined above, the empirical sibling covariance is an upper bound for family effects, since the covariance includes the effects of neighbourhood characteristics such as the quality and availability of community institutions and the characteristics of the other adults and children living in the neighbourhood.

Access to neighbourhood identifiers and family characteristics in the same data enables us to tighten the bounds mentioned above, both on the neighbourhood effect and on the family effects. Equation (3) suggets that the upper bound on the neighbourhood effects can be made tighter by introducing observed family characteristics shared by neighbours. Let $\tilde{\mathbf{X}}_{cf}$ denote such an observed subset of family characteristics. Following Solon *et al.* (2000) and Altonji (1988), we estimate a regression of the outcome variable on $\tilde{\mathbf{X}}_{cf}$, including a full set of neighbourhood dummy variables which will absorb the neighbourhood effects and the neighbourhood means of unobserved family characteristics. Let these within-neighbourhood estimates be denoted $\hat{\mathbf{a}}$. We expect the contribution from unobserved family characteristics to be positively correlated with our observed measure. Since we measure our family characteristics with error, and we cannot expect to fully proxy unobservables with observables, it seems reasonable to assume that $\text{cov}(\widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf}, \widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf'}) \leq \text{cov}(\widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf}, \widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf'})$. Our preferred measure is adjusted for observed family characteristics. We use the estimated neighbour covariance and subtract the covariance in predicted family effects,

$$\widehat{\text{cov}}_{adj.}(y_{cfs}, y_{cf's'}) = \widehat{\text{cov}}(y_{cfs}, y_{cf's'}) - \text{cov}(\widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf}, \widehat{\mathbf{a}}'\tilde{\mathbf{X}}_{cf'}). \tag{5}$$

Our assumptions on the positive sorting then implies that this measure is a tighter bound on the neighbourhood effects.

Even if all the relevant family characteristics were included in $\tilde{X}_{cf}$, and the associated parameter estimates were unbiased, the adjusted covariance would still be an upper bound since the covariance between the family effects and the neighbourhood effects is not accounted for. The linear additive form of (1) is an identifying assumption for the interpretation of the $\widehat{\text{cov}}_{adj.}(y_{cfs}, y_{cf's'})$ as an upper bound, as it restricts the potential for interaction between family and neighbourhoods.

Any upper bound on the variance of neighbourhood effects can be used to find a corresponding lower bound on the variance of family effects. By subtracting the adjusted neighbour covariance from the sibling correlation, what is left represents a lower bound on the variance of the family effects. The variance of the family effects can be written as

$$\text{var}(\alpha'\mathbf{X}_{cf}) = \text{cov}(y_{cfs}, y_{cfs'}) - \left[\text{cov}(y_{cfs}, y_{cf's'}) - \text{cov}(\alpha'\mathbf{X}_{cf}, \alpha'\mathbf{X}_{cf'})\right].$$

Since (5) is an upper bound on the two terms in brackets we use $\widehat{\text{cov}}(y_{cfs}, y_{cfs'}) - \widehat{\text{cov}}_{adj.}(y_{cfs}, y_{cf's'})$ as a lower bound on the family effects. The lower bound property arises from the fact that we cannot fully observe and correct for the tendency of similar families to cluster in the same neighbourhood. The covariance between the neighbourhood and the family effects, $\text{cov}(\alpha'\mathbf{X}_{cf}, \beta'\mathbf{Z}_c)$, does not influence the interpretation, since it is included in both the neighbour and the family covariance.

2. Data and Estimation

The database we use has been put together with sources from Statistics Norway (Moen *et al.*, 2003). It includes linked administrative data, which covers all people resident in Norway. We also have the national censuses of 1960 and 1970 (Vassenden, 1987). We can link records from these datasets using an unique personal identifier given to all Norwegian residents by the national population register. We use a set of household and census tract identifiers in the census to identify families and place of residence during childhood. The administrative data provide information about taxable income (excluding capital gains) and educational attainment, while we can use a variety of family background variables from the censuses.

2.1. *Neighbourhoods*

We use the individual's recorded census tract at the time of the census as an identifier of neighbourhood. Byfuglien and Langen (1983) document the principles used for delineating tract boundaries. In 1960 the main principles were that a 'densely populated area' with an expected population of at least 200 persons should be a separate tract, that tract boundaries should not cross parish boundaries, nor should they cross older administrative boundaries or boundaries that would result from expected adjustments of municipalities. Where population growth was expected, tracts should be planned such that adjustments of tracts in the following census would involve only a limited number of boundary adjustments. Finally, a tract should be homogeneous with respect to communications, industry and demographical structure. These regulations were not imposed on urban municipalities in 1960, and the size of urban tracts varies considerably in the 1960 census.

In 1970 the boundaries were redrawn to reflect changes in population density and a large number of municipality mergers during the 1960s. In 1960 the 732 municipalities had a total number of 7,996 tracts, while in 1970 the 451 municipalities had 8,818 tracts. The average tract populations were 464 and 439 in 1960 and 1970 respectively. In 1960, 6,127 tracts had a population of fewer than 500 individuals. This number grew to 6,809 in 1970.[4] Most of the new tracts appeared in urban areas, reflecting both urbanisation and that the formal guidelines for tract delineation only gradually were applied to urban areas. The tracts in Oslo, the capital city, had an average of 4,903 inhabitants in the 1960 census; this was reduced to 1,091 in 1970.

The Norwegian tracts were small by the international standards of the day. Sweden had 2,568 'parishes' in 1971, with an average of 3,145 individuals, Denmark had about 5,000 primary units in 1970, with an average of 990 individuals. Great Britain had 'enumeration districts' of about 750–1,000 individuals, in the 1961 census (Langen, 1975, pp. 5–6). The US Bureau of the Census requires the average population of all census tracts in a county to be about 4,000 people, and there were 62,276 tracts and Block Numbering Areas in the US 1990 census (Bureau of the Census, 1994, pp.10–1). The Norwegian census tracts are much closer in size to the US 'Block Groups', a subdivision of census tracts and block numbering areas.

The neighbourhood definitions used by Solon *et al.* (2000) are not census subdivisions. They use data from the Panel Study of Income Dynamics, and what makes it possible to identify neighbourhoods in the PSID is a strict hierarchical sampling procedure. Within each Primary Sampling Unit, smaller areas were chosen, such as 'cities, towns, census tracts etc.' (Solon *et al.*, 2000, p. 385). At least one 'chunk' of 20–30 contiguous dwellings was chosen from within each of these smaller areas, a total of 6–20 chunks per PSU. Within these chunks, 4 dwelling units were selected. From the information available, it seems reasonable to conclude that our neighbourhoods are somewhat larger than the neighbourhoods

[4] Langen (1975), Table 4.6 and and Table 4.7.

that can be identified from the PSID data, but smaller than the census tracts mostly used to assess neighbourhood effects using US data.[5]

We observe the neighbourhood children live in at one point in time. This may not accurately represent the environment of children from families who move around. Such measurement error will bias estimates of neighbourhood effects downward; however, families with children tend to move to neighbourhoods that are similar to those they leave, so we cannot conclude anything about the magnitude of this effect from statistics about the frequency of moving alone. One way to examine whether such moving introduces large biases is to compare the effects of 1960 neighbourhood on those who stayed to those who moved, and in Section 4 we examine this using a somewhat restricted and truncated sample that is different from our main sample: Because the tracts are not directly comparable across the two censuses, we construct aggregations of tracts that are comparable. Langen (1975, appendix D) provides a catalogue of 5,298 such comparable units. In many circumstances there were no changes made to tract boundaries, and the 'aggregation' consists of a single tract. But some of Langen's tract aggregations are very large, the largest being Oslo, the capital, in which the tracts were completely redrawn.[6] In order to examine how stable neighbourhood effects are, we will consider a subsample of Langen's aggregations. We restrict the sample to aggregations with fewer than 4,000 inhabitants in 1970, and exclude all tracts from the 1960 census that were split across municipalities in the years between 1960 and 1970. For the purpose of examining the effect of moving, we limit our sample to those aged 0–4 in the 1960 census, who we can expect to live with their parents at the time of the 1970 census.

2.2. *Outcome Variables and Observed Family Background*

We classify all children living in the same private household as siblings, excluding all institutional households. For parental classification, we use the recorded information on 'responsible adults' in the household. The 'responsible adults' are in the majority of cases biological parents, but to the extent that children live apart from their biological parents the head of the household and the spouse of the head are indicated as 'responsible adults' in the census. In 1960 only 1.5% of our sample lived in a household without any biological parents, compared to 3.2% in 1970.

Our measure of adult educational attainment of our main sample is taken from the register of the level of education maintained by Statistics Norway Vassenden (1995). This register provides a detailed code of the type of the highest completed education, the completion date and how many years of schooling the highest completed education corresponds to. For individuals with no recent education, their level of education as of the 1970 census is recorded.

[5] Studies such as Kremer (1997), Topa (2001) and Conley and Topa (2002) are examples using US census tracts as local neighbourhoods, but in different frameworks than ours. Solon *et al.* (2000 endnote 9) note that the average size of lowest-level units in the National Longitudinal Study of Youth is 200–250 dwelling units.

[6] As the research leading to these aggregations was financed by a programme on rural regions, the lists linking addresses to tracts in urban areas were not used.

Our measure of adult earnings is constructed from administrative data that are collected from tax returns and various government agencies. We use the 1990–5 observations of a category of earnings that is used to calculate accumulation of insurance benefits. This definition includes wages, income from self-employment, unemployment benefits and sick-leave payments, but excludes capital income, social assistance, pensions and other transfers. Unemployment insurance and sickness benefits are included as these are conditional on previous employment and we do not want fluctuations due to transitory unemployment or sickness periods to affect our estimates. We inflate all numbers by the Consumer Price Index, and exclude all observations from before the completion of education or of less than NOK 10,000 at 1995 prices. Since there may be secular trends and life-cycle effects in our outcome measures, all the numbers we calculate in this article are residuals from a regression on dummy variables indicating the year of birth, following Solon *et al.* (2000). We calculate the mean of the logarithm of these observations for each individual.

Whereas our measure of adult education is the natural one, our measure of adult earnings is potentially more problematic (Haider and Solon, 2004). The parameters we estimate are only defined for those who actually participate in the labour market and patterns in measurement error caused by differential participation rates may bias our results. However, we only exclude a small proportion of the population because of restrictions on earnings. Table 1 shows that only about 2.5% of males and 6.5% of the females have earnings below the threshold in all six years. More than 75% of males have earnings above the threshold in all six years, this proportion is somewhat lower for women. Since participation is high and fairly constant for both genders, exclusion of observations below the threshold is unlikely to cause major biases.

Even if differential participation does not cause large systematic biases, it could still be the case that our measure is very poorly related to lifetime labour earnings. We are unable address this question for the sample we use in this article. We have, however, looked into how our measure performs over the life-cycle of the 1942 cohort, for which we have labour earnings (same definition used in this article) for 1967–2002. This covers the ages 25–60, for which we have calculated an 'ideal'

Table 1

The Distribution of Years of Earnings by Gender and Birth Cohorts

	Years behind earnings measure						
	0	1	2	3	4	5	6
Males							
1946–55	0.028	0.023	0.027	0.024	0.037	0.040	0.820
1956–65	0.023	0.035	0.042	0.039	0.051	0.060	0.751
Females							
1946–55	0.067	0.040	0.040	0.039	0.060	0.053	0.703
1956–65	0.065	0.053	0.055	0.055	0.069	0.079	0.625

Note: Numbers of years 1990–95 in which earnings are above NOK (1995) 10,000. Based on the sample of all residents in the two age-groups that are present in the relevant censuses.

Table 2

Comparison of Samples with Population

	1946–55 cohorts			1956–65 cohorts		
	Earnings	Education	Pop.	Earnings	Education	Pop.
Male:						
mean age	44.48	44.48	44.56	34.38	34.39	34.41
mean education (years)	11.53	11.45	11.58	11.89	11.86	11.91
(standard deviation)	2.89	2.92	2.93	2.30	2.39	2.30
mean log av earnings 1995	12.31	12.30	12.31	12.19	12.20	12.20
(standard deviation)	0.57	0.58	0.58	0.55	0.55	0.55
share full time working 1995	0.735	0.704	0.714	0.733	0.720	0.726
share unemployed 1995	0.075	0.074	0.072	0.108	0.108	0.107
number of people	106 287	113 739	290 345	122 413	125 436	297 734
Female:						
mean age	44.49	44.48	44.55	34.40	34.41	34.47
mean education (years)	11.00	10.88	10.99	11.84	11.74	11.82
(standard deviation)	2.60	2.60	2.60	2.29	2.27	2.28
mean log av earnings 1995	11.77	11.76	11.77	11.67	11.66	11.68
(standard deviation)	0.62	0.63	0.63	0.66	0.66	0.66
share full time working 1995	0.460	0.429	0.430	0.436	0.408	0.412
share unemployed 1995	0.064	0.064	0.063	0.102	0.103	0.100
number of people	92 581	103 109	278 381	103 308	114 549	286 074

Note: The table compares the samples that remains after linking the administrative files with the census files and restricting the sample to tracts with at least two families of two individuals with non-missing observations.

measure: present value of life-time labour earnings. We then calculate the measure we use in this article at the relevant age-groups and correlate this with the logarithm of present value of labour earnings. For the age-groups with comparable participation rates we then have correlations between log present value of earnings and our measure between 0.77 and 0.82.[7] These correlation suggest that our measure of earnings approximates life-time earnings quite well. While there might well be some attenuation bias from measurement error, it is unlikely that systematic changes in attenuation invalidates our analysis of changes in sibling and neighbour correlations.

Table 2 provides summary statistics of our sample compared to the full population from the administrative data in 1995. To be included in our sample, a family must include at least two children aged 5–14 at the time of the census. We restrict ourselves to the families with at least two brothers or two sisters in the relevant age-group. This restriction implies that our estimating samples are smaller than the full population of the relevant age-group. Even so, it does not seem that the

[7] The correlation is lower for the younger women, at 0.62, but participation rates for young women of the 1942 cohort are also very low and different from those in Table 1: 27% have no years of earnings and only 28% have earnings above the threshold in all years.

samples are much different from the full population in terms of observed characteristics. There is a limited increase in the average years of education from the older to the younger cohorts, and this increase is stronger for women. The variance of the years of educations decreases from the older to the younger cohorts. The same is true of annual earnings, but this may simply reflect that earnings are measured at two different stages of the life-cycle. Our samples are orders of magnitude larger than that of Solon *et al.* (2000), who use 687 individuals from 144 clusters to examine educational attainment, Page and Solon (2003) use 443 individuals from 120 clusters to examine male earnings.

The measure we have of the educational attainment of the *parents* of individuals from our main sample is different. From the 1970 census we have education recorded as years of schooling. The 1960 measure of education is a categorical classification. We could in principle use a set of dummy indicators to correct for parental education but this would identify the coefficients from those neighbourhoods with two or more parents with the same educational background, and the 1960 and 1970 measures would not be comparable. We have therefore transformed the categorical parental education codes into years of education. We use repeated observations of the same individuals to construct a mapping from the 1960 codes to years of schooling, and then apply this mapping to all parents, regardless of whether we have in fact repeated observations of this parent.

2.3. *Estimation*

Estimation of the covariance of some characteristic within a group is not a difficult problem. There are many ways to combine these within-group estimates, but note that observations here consist of pairs of siblings. A family of 2 siblings contributes one such pair, a family of 3 contributes 3 and so on: With S siblings, there are $S(S-1)/2$ unique pairs. Solon *et al.* (2000) provides (6) as an estimator of the covariance between siblings of a variable y with $\mathrm{E}(y) = 0$,

$$\sum_{c=1}^{C} W_c \left\{ \sum_{f=1}^{F_c} W_{cf} \left[\sum_{s \neq s'} \frac{y_{cfs} y_{cfs'}}{S_{cf}(S_{cf}-1)/2} \right] \Bigg/ \sum_{f=1}^{F_c} W_{cf} \right\} \Bigg/ \sum_{c=1}^{C} W_c. \tag{6}$$

Here c denotes neighbourhood, f denotes family and s denotes sibling, the W_c and W_{cf}s are weights and S_{cf} is the number of siblings in family f in neighbourhood c. Solon *et al.* considered four different weighting schemes. In practice, we have found the differences among estimates with different weighting schemes to be negligible. All estimates in this article give all sibling-pairs and neighbour-pairs equal weight regardless of whether they came from large or small families and neighbourhoods. To centre the observation on zero, we follow Solon *et al.* and first regress the variable in question on dummies for each year of birth.

The estimation of neighbourhood correlations is complicated by the fact that we want the correlation of one individual in a family with all other individuals except its siblings, so that the neighbourhood covariance is not contaminated by sibling correlations in small neighbourhoods. For a pair of families with S_{cf} and $S_{cf'}$ siblings there are $S_{cf}S_{cf'}$ unique pairs, and if there are F families in the neighbourhood, there are $F(F-1)/2$ unique family pairs. Solon *et al.* provide an estimator of the neighbourhood correlations that is similar in spirit to (6),

$$\sum_{c=1}^{C} W_c \left[\sum_{f \neq f'} W_{cff'} \left(\sum_{s=1}^{S_{cf}} \sum_{s'=1}^{S_{cf'}} \frac{y_{cfs} y_{cf's'}}{S_{cf} S_{cf'}} \right) \Big/ \sum_{f \neq f'} W_{cff'} \right] \Big/ \sum_{c=1}^{C} W_c. \tag{7}$$

In order to adjust for the effect of observed family characteristics, we follow Solon *et al.* (2000) and regress y_{cfs} on a vector of observed characteristics $\tilde{\mathbf{X}}_{cf}$ and neighbourhood dummies. Subtracting the covariance of the predicted family-effects from the total covariance, as indicated in (5), and dividing by the total variance of y_{cfs}, we obtain neighbour correlations that are adjusted for observed family characteristics.

We have no analytical expression for the variance of these estimators. Solon *et al.* use balanced half-samples to take into account the complex sampling procedure of the PSID data. We have chosen a bootstrap procedure that resamples blocks of data and repeats all initial corrections for every resampled set of data. Since neighbourhoods close together may well be have spatially correllated outcomes, we use municipalities as units of resampling for both the sibling and neighbour correlations. Since we have complete coverage of a whole economy that is spread over a relatively large geographical area, we believe that there is limited room for spatial correlation to bias the standard errors, and that our strategy of resampling on municipalities provides us conservatively large estimates of the standard errors. Calculation of the bootstraps are computationally expensive, and we have limited the number of replication to 250.[8]

3. Results

In this Section we first report the empirical sibling and neighbour correlations. Then we present our preferred estimates of the upper bounds of neighbourhood effects that accounts for sorting on observed family characteristics. Finally, we report family effects, net of the impact of siblings growing up in the same neighbourhoods. In all cases, we show the estimates by gender, sets of birth cohorts and adult outcome measure. Several robustness checks are performed to back our interpretation of the numbers as reliable estimates of neighbourhood and family effects, including discussions of how regional differences and neighbourhood misclassification affect our estimated neighbour correlations. We close this section with a discussion of policy changes during the 1950s, 1960s and 1970s that are candidates for explaining our results.

3.1. *Sibling and Neighbour Correlations in Education and Earnings*

The sibling and neighbour correlations are shown in Table 3. The brother correlation in years of schooling is around 0.42 and somewhat higher, around 0.46,

[8] With 250 replications there is still some variation in the standard errors when starting at different random seeds, but this variation is much to small to influence inference in any important way (evidence available on request).

Table 3

Correlation in Education and Adult Earnings Among Siblings and Neighbouring Children.

	Siblings		Neighbours	
	1946–55	1956–65	1946–55	1956–65
Education 1995				
Males	0.4150	0.4213	0.1121	0.0612
	(0.0088)	(0.0075)	(0.0261)	(0.0075)
Adjusted for parental education (PE)			0.0590	0.0245
			(0.0111)	(0.0030)
Adjusted for family structure (FS)			0.1105	0.0602
			(0.0260)	(0.0076)
Adjusted for PE and FS			0.0494	0.0206
			(0.0094)	(0.0032)
Adjusted for PE, FS and parental income				0.0163
				(0.0034)
Females	0.4561	0.4739	0.1027	0.0653
	(0.0064)	(0.0080)	(0.0213)	(0.0095)
Adjusted for parental education (PE)			0.0493	0.0245
			(0.0062)	(0.0046)
Adjusted for family structure (FS)			0.1013	0.0642
			(0.0046)	(0.0050)
Adjusted for PE and FS			0.0405	0.0202
			(0.0205)	(0.0095)
Adjusted for PE, FS and parental income				0.0153
				(0.0050)
Average log Earnings 1990–95				
Males	0.2032	0.1845	0.0591	0.0283
	(0.0082)	(0.0059)	(0.0074)	(0.0051)
Adjusted for parental education (PE)			0.0499	0.0252
			(0.0059)	(0.0047)
Adjusted for family structure (FS)			0.0584	0.0280
			(0.0077)	(0.0048)
Adjusted for PE and FS			0.0470	0.0245
			(0.0053)	(0.0052)
Adjusted for PE, FS and parental income				0.0221
				(0.0051)
Females	0.1480	0.1645	0.0292	0.0201
	(0.0053)	(0.0043)	(0.0055)	(0.0024)
Adjusted for parental education (PE)			0.0225	0.0141
			(0.0041)	(0.0022)
Adjusted for family structure (FS)			0.0287	0.0197
			(0.0051)	(0.0024)
Adjusted for PE and FS			0.0206	0.0127
			(0.0036)	(0.0021)
Adjusted for PE, FS and parental income				0.0104
				(0.0021)

Note. Estimated on the full population of those aged between 5–15 in the year of the census in families with at least 2 children in this age span, and in neighbourhoods with at least two such families. Correlations in education based on 1995 data, the earnings measure is the mean of logarithm of earnings 1990–5, dropping those years before completion of education or with less than NOK (1995) 10,000 in earnings.

The variables for parental education include a 4th degree polynomial in mother's and father's education and a first degree interaction term and dummies for whether any of these are missing. The family structure variables include indictors for whether parents are currently divorced, separated, presence of a step-parent or non-biological parents, single parents and size of the household. Parental income consists of the logarithm of the income of the responsible adults and indiciators for whether any of these are missing.

for sisters. These figures are surprisingly similar to those found for other countries. According to Solon (1999), sibling correlations in years of education in the US are a little higher than 0.5. Table 3 also reveals that the gender-specific education correlation is fairly stable across cohorts, with a slightly higher correlation for the younger birth cohorts. Thus, the total impact of family and neighbourhood characteristics shared by siblings seems to be constant over time. The neighbour correlations in education are much lower. While the male correlation is higher among those born 1946–55, the gender difference is negligible in the younger cohorts. While Solon *et al.* (2000) report an unadjusted neighbour correlation of around 0.2 in the US, neighbourhoods seem to be less important in Norway.[9] Moreover, the impact of location during childhood seems to have fallen over time, since the neighbour correlation is considerably lower in the younger birth cohorts. It drops from 0.112 to 0.061 for males and from 0.103 to 0.065 for females.

Correlations are considerably lower when we look at earnings. The sibling correlations in the range of 0.15–0.20 are similar to figures found in previous studies. Björklund *et al.* (2002) find that brother correlations in Nordic countries are significantly lower than the typical 0.3–0.4 found in recent US studies. In spite of the higher education correlation for sisters, we find that earnings are less correlated among sisters than among brothers.[10] The brother correlation drops from the oldest to the youngest cohorts, from 0.203 to 0.185, while the earnings correlation for sisters increases from 0.148 to 0.165.

The neighbour correlations in adult earnings are positive, and higher for neighbouring boys than for girls. For the older birth cohorts, the correlation in adult earnings among neighbouring boys is 0.059 and 0.029 for girls. The higher male correlation is found in both cohorts. As for education, the male correlation is reduced by approximately one half from the 1946–55 to the 1956–65 cohorts and somewhat less for females.

To summarise, the correlations indicate that both families and neighbourhoods matter. Local childhood communities have become less important in explaining the overall variation in adult outcomes over time, and the gender difference has been reduced as well.

3.2. *Neighbourhood Effects*

The neighbour correlations in Table 3 are upward biased measures of the true influence neighbourhoods have on individuals, due to sorting of families into communities. Resemblance in adult outcomes among persons who spent their childhoods in the same community may partly, or even completely, reflect that neighbouring children experience similar family environments. In Table 3 we also report adjusted neighbour correlations in years of schooling where we subtract the covariance in effects of observed family characteristics.[11]

[9] Note, however, that the estimates of Solon *et al.* (2000) have standard errors around 0.05.

[10] Bound *et al.* (1986) finds that sisters are *more* similar than brothers, but they examine residuals from wage equations.

[11] To prevent family background coefficients from being affected by neighbourhood effects, these are estimated with neighbourhood dummies.

When we partial out the effect of parental education the correlations drop considerably. Family structure characteristics are less important, as the numbers are very similar to the unadjusted figures. When combining the two sets of family background variables available for both cohorts, the correlation drops considerably, from 0.112 to 0.049 for males and from 0.103 to 0.041 for females in the 1946–55 cohorts. The impact of the family background adjustment is similar for the 1956–65 cohorts. The neighbour correlations for males and females have become more similar over time. Parental income information is only available for the younger cohort and this adjustment reduces the correlation slightly.

In Table 3 we also report bounds on the neighbourhood effects on adult earnings. After subtracting the contribution of parental education, we find that the neighbour correlations fall in both cohorts and for both men and women. Family structure is less important for education. Including parental income among the family controls has a negative, but modest, effect on the correlations.

Several important conclusions can be drawn from the estimates in Table 3. First, observed family sorting into neighbourhoods does not fully explain the resemblance in adult earnings among persons who grew up in the same neighbourhood. Second, neighbourhoods have become less important as determinants of adult earnings. This is consistent with, and presumably partly explained by, the declining effects of neighbourhoods on educational attainment. Third, childhood location seems to have stronger effects on adult earnings for males than for females. This gender difference in earnings correlations among neighbouring children is present for both cohorts. Finally, the family background adjustment wipes out the difference between earnings and schooling correlations. While the unadjusted neighbour correlations are much higher for education than earnings, they are strikingly similar after having adjusted for observed family sorting into neighbourhoods, especially for the younger cohorts.

One might ask what the correlations mean in terms of absolute size of neighbourhood effects. If we consider male earnings, we use the upper bound and the variance from Table 2, and find that the standard deviation of the neighbourhood effects to be smaller than 0.124 and 0.086 for the older and younger cohorts respectively.[12] Although the correlations may seem low, the variance is large enough that the bounds on neighbourhood effects are in no way negligible.

3.3. *Family Effects*

Previous studies of sibling correlations do not disentangle family from neighbourhood effects. The only exception, to our knowledge, is Page and Solon (2003). Table 3 suggests that family background is by far more important than neighbourhoods, but the sibling correlations also include the effects of siblings growing up in the same neighbourhoods. In this Section we tighten the bound on family effects using the method described in Section 1.

The lower bound on family effects on education is 0.366 for older males. In the younger cohorts, we find that at least 0.401 of the male variance in schooling is

[12] The estimates of neighbour correlation are those adjusted for parental education and family structure. The bound is calculated as $\mathrm{sd}(\boldsymbol{\beta}'\mathbf{Z}_c) \leq \mathrm{sd}(y_{cfs})\sqrt{\mathrm{corr}_{adj.}(y_{cfs}, y_{cf's'})}$.

explained by factors, other than childhood location, which siblings share. The female estimates are 0.416 and 0.454, respectively. Thus the lower bound on family effects has increased slightly for both men and women.[13]

Considering male earnings, the lower bound on family effects is very similar for the two cohorts, 0.156 among those born 1946–55 compared to 0.160 for those born ten years later. Consequently, weaker neighbourhood effects could potentially account for all of the decrease in brother correlations in Table 3.

For females, the lower bound on the family effects on adult earnings is actually increasing from 0.127 in the older cohorts compared to 0.152 in the younger. This suggests that the higher earnings correlation among brothers compared to sisters in the older cohorts is partly explained by larger neighbourhood effects on earnings for males. While family effects on educational attainment seem to be somewhat stronger for females than for males, the gender difference is smaller for earnings. The lower bound on family effects on earnings in the younger cohort is basically the same for both genders.

Since it seems that the impact of neighbourhoods has declined, one might wonder why the sibling-correlation did not decline correspondingly. But such a decline a is not a necessary consequence. If fewer families are financially constrained when making human investments, other mechanisms which tend to create homogeneous families may well become more important. For instance, if ability sorting into education is important, it may well be that poorer families, formerly restricted from investing in all of their promising children, can afford to send all of them on to higher education as the effective price of attendance drops.

3.4. *Robustness Checks*

In this Section we discuss whether the neighbour correlations, and particularly the drop over time, could reflect other mechanisms than declining neighbourhood effects. We first check whether the neighbourhood effects represent permanent regional effects rather than impact of childhood local community. Then we address the decline in light of changes in neighbourhood boundaries, sorting, measurement error in permanent earnings and misclassification of childhood community. These are all alternative explanations for why the upper bound on neighbourhood effects is reduced.

The neighbour correlations may reflect impact of regional characteristics rather than local effects. In other words, what we label neighbourhood effects may represent permanent differences between children growing up in different regions (e.g. urban vs non-urban areas). Moreover, in the case of adult earnings it may also reflect interaction of location preferences and regional price levels (Griliches, 1979). It is well documented that workers in urban areas in the US are paid a premium for living and working there.[14] If resemblance in adult earnings among

[13] All the bounds discussed in this Section are based on the largest set of family background characteristics that is available for both the older and younger cohorts.

[14] Using our measure of earnings in a Mincer earnings equation including schooling, experience, experience squared, gender and regional dummies, the largest difference is 0.125 log points between the counties Aust-Agder (with low earnings) and Akershus.

neighbouring children simply reflects the interaction between geographical location preferences and regional wage and price differences, heavily influenced by where they grew up, it is only weakly related to what people think of as neighbourhood effects on adult welfare. Since we neither have a model of geographical mobility, nor good regional wage and price indices, we perform a simple check where neighbour correlations are estimated within childhood county. We add 20 county dummies to birth year variables in the process of constructing the outcome measure, and we only predict family effects on within county variation. In this way we condition out all effects of adult location that are explained by the region in which the person spent his childhood and thereby all variation in neighbourhoods across counties. If all the effects are associated with region rather than neighbourhood, or if location preferences explained the correlations, we would expect the neighbour correlations to vanish. Table 4 reports the sibling and neighbour

Table 4

Correlation in Education and Adult Earnings Among Siblings and Neighbouring Children. Within Regions

	Siblings		Neighbours	
	1946–55	1956–65	1946–55	1956–65
Education 1995				
Males				
Within childhood county	0.4057 (0.0070)	0.4177 (0.0067)	0.0885 (0.0182)	0.0562 (0.0065)
Adjusted for parental education (PE)			0.0587 (0.0121)	0.0270 (0.0032)
Adjusted for family structure (FS)			0.0879 (0.0184)	0.0556 (0.0061)
Adjusted for PE and FS			0.0563 (0.0135)	0.0247 (0.0031)
Adjusted for PE, FS and parental income				0.0220 (0.0035)
Females				
Within childhood county	0.4495 (0.0052)	0.4694 (0.0070)	0.0787 (0.0112)	0.0622 (0.0082)
Adjusted for parental education (PE)			0.0485 (0.0051)	0.0295 (0.0049)
Adjusted for Family structure (FS)			0.0781 (0.0114)	0.0616 (0.0086)
Adjusted for PE and FS			0.0460 (0.0058)	0.0270 (0.0043)
Adjusted for PE, FS and parental income				0.0240 (0.0049)
Average log Earnings 1990–95				
Males				
Within childhood county	0.1831 (0.0065)	0.1714 (0.0052)	0.0310 (0.0036)	0.0131 (0.0023)
Adjusted for parental education (PE)			0.0259 (0.0029)	0.0131 (0.0020)
Adjusted for family structure (FS)			0.0307 (0.0036)	0.0107 (0.0020)
Adjusted for PE and FS			0.0251 (0.0033)	0.0129 (0.0022)
Adjusted for PE, FS and parental income				0.0087 (0.0022)

Table 4

Continued

	Siblings		Neighbours	
	1946–55	1956–65	1946–55	1956–65
Females				
Within childhood county	0.1406	0.1579	0.0127	0.0117
	(0.0045)	(0.0046)	(0.0013)	(0.0012)
Adjusted for parental education (PE)			0.0089	0.0071
			(0.0014)	(0.0012)
Adjusted for family structure (FS)			0.0125	0.0115
			(0.0013)	(0.0012)
Adjusted for PE and FS			0.0083	0.0063
			(0.0014)	(0.0012)
Adjusted for PE, FS and parental income				0.0050
				(0.0013)

Note. Estimated on the full population of those aged between 5–15 in the year of the census in families with at least 2 children in this age span, and in neighbourhoods with at least two such families. Correlations in education based on 1995 data, the earnings measure is the mean of logarithm of earnings 1990–5, dropping those years before completion of education or with less than NOK (1995) 10,000 in earnings. Both the outcome variable and the observed family characteristics have their county mean subtracted.
The variables for parental education include a 4th degree polynomial in mother's and father's education and a first degree interaction term and dummies for whether any of these are missing. The family structure variables include indictors for whether parents are currently divorced, separated, presence of a step-parent or non-biological parents, single parents and size of the household. Parental income consists of the logarithm of the income of the responsible adults and indiciators for whether any of these are missing.

correlations, estimated within childhood county. By comparing Table 3 and Table 4, we find that the brother and sister correlations in adult outcomes are basically unchanged.

The neighbourhood effects on educational attainment drop significantly, but do not disappear. Considering schooling years of the oldest cohorts first, the neighbour correlations are reduced by approximately 25% when we estimate within regions. Adjusting for childhood county has no important impact on the younger cohorts, although the within-county estimates are lower for both men and women.

Earnings correlation estimates are more affected by the childhood-region adjustment. This is what we would expect if they were driven by regional wage and price differentials, since these will partly be accounted by adult location being strongly affected by childhood region. Comparing Tables 3 and 4 we see that the neighbour correlations in earnings drop by about 50% when we estimate within counties. We also report the family background adjusted estimates in Table 4 following the same procedure as in Table 3. The adjustment does reduce the correlations, especially for females where the neighbour correlations in earnings end up below 0.01. The lower effects of neighbourhoods in the younger cohorts remain for both males and females. It is important to emphasise that the within-county estimates neglect all the effects of variation in neighbourhoods between counties and this downward bias can be substantial. To summarise, neighbourhood effects remain positive and declining even if we estimate within counties.

The next question is whether the change in neighbourhood boundaries between 1960 and 1970 is responsible. However, neighbourhoods were more narrowly defined in 1970 and the average number of residents was lower than in 1960. As the 1970 classification represents smaller communities, one would expect the estimated correlations to increase rather than fall.

In the cohort comparison we measure income at different points of the life cycle. The two sets of cohorts are measured while 25–40 and 35–50 years of age respectively. Whereas we have preferred to use permanent earnings, the correlations will be biased downward by measurement error or transitory components in our measure of average earnings (Solon *et al.*, 1991). One might expect that we have a more noisy earnings measure for the younger cohort, potentially explaining the drop in the neighbour correlation. But any important bias of this kind would also affect the sibling correlations. With a simple measurement error model, the estimated correlation is proportional to the correlation in permanent earnings. Assume that any difference between our observed earnings and permanent earnings is generated by the same model for siblings and neighbours. If the drop in the estimated neighbour correlation was generated by a different degree measurement error, we would expect a similar relative decline in the sibling correlations. The stability, and even slight increase for women, suggests that this bias is likely to be small.

We interpret the estimates as upper bounds on the neighbour effects and the true effects can be constant if sorting on unobserved family characteristics has decreased over time. We check this explanation by looking at how adult education is distributed within and between neighbourhoods over time, since we expect any trend in sorting on unobserved family background to be similar to the trend in sorting on observed characteristics. If sorting decreased from 1960 to 1970, one would expect the between-neighbourhood component of the total variance in adult education to fall. Table 5 shows the opposite. Consequently, weaker sorting leading to an upper bound which is closer to the true neighbourhood effect in the younger cohorts does not seem to be the explanation.

Finally, community misclassification error may explain the drop in neighbourhood correlation if family mobility increased during the 1950s and throughout the 1960s. We only observe location at one point in time. Since families move,

Table 5

Degree of Neighbourhood Sorting

	Mother's education		Father's education	
	1960	1970	1960	1970
Mean	8.008	8.627	8.641	0.932
$\hat{\sigma}_u$	0.557	0.900	0.935	1.269
$\hat{\sigma}_\epsilon$	1.578	1.194	2.274	2.516
$\hat{\rho} = \hat{\sigma}_u^2/(\hat{\sigma}_u^2 + \hat{\sigma}_\epsilon^2)$	0.104	0.177	0.145	0.203

Note. Decomposition of the variance of adult educational attainment at the time of the censuses. Estimates from the regression $E_{ic} = \bar{E} + u_c + \epsilon_{ic}$ (with neighbourhood fixed effects). Calculated on the population of individuals aged 30–50 at the time of the censuses.

neighbourhood effects tend to be downward biased for both cohorts. On the other hand, families tend to move to similar neighbourhoods. A recent study for the US (Kunz *et al.*, 2003) shows that one-year observation of location does not create a significant bias in the neighbourhood effects, but this conclusion may not be applicable to Norwegian data. This misclassification explanation cannot be checked for the full samples, as changing boundaries makes a substantial number of neighbourhoods non-comparable over time. Using the subset of comparable geographical units described in Section 1, we estimate different correlations of those who stay in the same and those who move to a different neighbourhood during the 1970s. We limit ourselves to those aged 0–4 in 1960, because we want to follow children who can be expected to live with their parents ten years later. As we discussed in Section 1, these tract-aggregates contain more individuals and they cover larger geographical areas than the original units.

The first column in Table 6 report the neighbour correlations for the 1956–60 cohorts, based on those individuals who spent at least some of their first five years in the aggregated tracts, but with the same definitions of neighbourhoods as in the previous Tables. The second column is based on the aggregated tracts, using the same set of individuals. As expected, aggregation implies that correlations fall somewhat. Splitting stayers and movers in column three and four, all correlations are lower for movers than for stayers, suggesting that mobility does create a slight downward bias. For male earnings, the correlation among stayers is twice that of movers. Even if misclassification may have increased over time, the magnitude of the difference between stayers and movers is too small to provide an explanation for the drop in neighbour correlations over time.

Since changing neighbourhood boundaries, weaker sorting of families into neighbourhoods, measurement error in earnings or misclassification of neighbourhoods cannot explain the drop in neighbour correlations over time, we are fairly confident that the true impact of location during childhood has declined.

Table 6

Stayers and Movers

	All (tracts)	All (aggregations)	Stayers	Movers
Education				
Male	0.0593	0.0444	0.0482	0.0363
	(0.0074)	(0.0124)	(0.0130)	(0.0119)
Female	0.0589	0.0400	0.0418	0.0311
	(0.0102)	(0.0156)	(0.0160)	(0.0131)
Earnings				
Male	0.0334	0.0296	0.0377	0.0173
	(0.0031)	(0.0056)	(0.0057)	(0.0052)
Female	0.0201	0.0155	0.0164	0.0114
	(0.0023)	(0.0035)	(0.0035)	(0.0028)

Note. Neighbour correlations for sample of stayers and movers. The sample includes those aged 0–5 in 1960 and with mothers living in one of 4,969 tract aggregations with fewer than 4,000 inhibitants and not containing 1960 tracts that were split among several tracts in the 1970 census. The first column summarises this sample at the lower tract level, stayers and movers inclusive. The group "stayers" lived in the same tract aggregation in 1970 as in 1960, while the "movers" (27.5% of the sample) had moved to some different tract aggregation.

Although attributing this change to public policies must be speculative, we want to point to some reforms and policy changes that we find likely to have played a role.

3.5. *Possible Explanations of Declining Neighbourhood Effects*

Local government services have been an important component in the building of the Norwegian welfare state after the Second World War. In the late 1940s real per capita local government spending increased by an annual rate of 9%, remaining at a high level of around 5% during the next three decades Borge and Rattsø, 2002). As a consequence, local public spending as a percentage of GDP increased from 9% in the late 1940s to around 16% in the 1970s. During the same period, the relative variation in spending across municipalities declined sharply. In the years before and after the Second World War, the tax base given by the local private income level largely explained the variation in spending across municipalities. Redistributive measures such as central government grants to municipalities were gradually introduced, and by 1980 the correlation between current per capita municipal spending and private income had changed from large and positive to negative. As far as neighbourhood institutions providing primary school and health care services are concerned, the first three decades after the Second World War were characterised by an overall expansion and an equalisation of spending across municipalities (Falch and Tovmo, 2003).

The school reforms implemented in the 1960s and 1970s are particularly interesting as possible explanations for the drop in neighbourhood effect for the cohorts growing up around 1960 and 1970. Norway experienced a sequence of school reforms during this period. The reform of the primary school system during the 1950s introduced a common curriculum in all communities, as well as access to the same number of teaching hours throughout the country. The comprehensive school reform of the 1960s increased the minimum level of schooling from 7 to 9 years, unified the education system and provided a common curriculum for all schools. In the 1970s regional colleges were established to enhance equality of opportunity in terms of transition to higher education for people growing up in all regions. The total number of students in higher education grew by 53% between 1971 and 1981 (Statistics Norway, 2001, Table 190).

The comprehensive school reform that was implemented between 1960 and 1970 was probably the most extensive. This aimed at raising the level of education, smoothing the transition into higher education and enhancing equality of opportunities across socio-economic and geographical backgrounds. It is expected that this reform reduced the effect of family background as well as neighbourhoods from 1960 to 1970. While most of the cohorts born between 1946–55 completed compulsory education before the reform, the 1956–65 cohorts went to the new comprehensive schools. Analyses of the participation rate to higher education for cohorts born from 1942–70 show a strong degree of regional equalisation (Haegeland *et al.*, 1999). The comprehensive school reform weakened the impact of socio-economic background on transition to higher education (Aakvik *et al.*, 2003).

Access to student grants and loans was expanded in the late 1960s and early 1970s. A grant for students older than 16 who lived more than 40 km away from their parents was introduced in 1968. Generally, from the age of 18 all students were entitled to a subsidised loan which covered living expenses. Tuition fees at Norwegian universities have been negligible. One motivation for the student grant and loan scheme introduced during the late 1960s was to promote equality of opportunity, such that educational qualifications could be attained independent of geographical location, age, gender, economic or social status.[15]

4. Concluding Remarks

Family background and childhood neighbourhood play an important role in explaining adult education and earnings. While most studies evaluating the combined effects of family and neighbourhoods are from the US, we present evidence from a country with different labour market institutions, educational system and welfare policies. Census data from Norway enable us to construct neighbourhoods and use a detailed set of family background variables. We focus on whether the impacts of neighbourhoods and families have changed over time, estimating separate parameters for children born 1946–55 and those born ten years later. Our main results can be summarised as follows.

Neighbour correlations in years of schooling for the 1946–55 birth cohorts are 0.112 for boys and 0.103 for girls. The log earnings correlations are estimated to be 0.059 and 0.029, respectively. Comparing the 1946–55 with the 1956–65 birth cohorts, we find a declining effect of neighbourhoods as the correlations are reduced by approximately one half.

Neighbour correlations are upward biased estimates of the influence of local childhood environments because similar families cluster in communities. We tighten the bound on the variance of neighbourhood effects by using data on observed family background. Adjusting for observed family background, the correlations drop considerably, for education down to 0.043 and 0.041 for the oldest boys and girls, respectively. Earnings correlations among neighbouring children born 1946–55 are reduced to 0.047 and 0.021, for boys and girls respectively. In the younger cohorts the neighbour correlations are about half that of the older cohorts.

We discuss whether the decline in neighbour correlation reflects changes in neighbourhood boundaries or reduced sorting of families into communities, and reject these explanations. Although higher geographical mobility among parents and measurement error in adult earnings for the younger cohorts *may* contribute to lower earnings correlations among neighbouring children over time, these factors only explain a minor part.

The impact of families, net of neighbourhood effects shared by siblings, is found to be fairly stable across cohorts. For adult earnings we find adjusted brother

[15] This was later formulated in the first paragraph in "Lov om utdanningsstøtte til elever og studenter", law of 26.04.1985, no. 21.

correlations of 0.156 and 0.160, for the older and younger cohorts respectively. The corresponding sister correlations are 0.127 and 0.152, suggesting a convergence between genders.

Families and neighbourhoods have weaker effects on adult outcomes in Norway than in the US, adding to the evidence that intergenerational mobility is higher in the Scandinavian welfare states than in the US (Björklund and Jäntti, 1997). We offer no rigorous tests of why neighbourhoods explain a lower fraction of the variation in adult outcomes among the younger cohorts; however, we single out the expansion of local government services in general, and education reforms in particular, as important candidates. These policies were implicitly targeted to promote equality of opportunity. These policy reforms affected the post-war birth cohorts differently. In particular, those born after 1955 faced a more similar school system and lower costs of educational investment than those born during the previous decade.

Ragnar Frisch Centre for Economic Research
Norwegian School of Economics and Business Administration
Norwegian School of Economics and Business Administration

Submitted: 22 January 2002
Accepted: 28 February 2005

References

Aakvik, A., Salvanes, K. G. and Vaage, K. (2003). 'Measuring heterogeneity in the returns to education in Norway using educational reforms', IZA Discussion Paper 815.

Altonji, J. G. (1988). 'The effects of family background and school characteristics on education and labor market outcomes of relatives', mimeo, Northwestern University.

Björklund, A., Eriksson, T., Jäntti, M., Raaum, O. and Österbacka, E. (2002). 'Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States', *Journal of Population Economics*, vol. 15(4), pp. 757–72.

Björklund, A. and Jäntti, M. (1997). 'Intergenerational income mobility in Sweden compared to the United States', *American Economic Review*, vol. 87(5), pp. 1009–18.

Black, S. E. (1999). 'Do better schools matter? Parental valuation of elementary education', *Quarterly Journal of Economics*, vol. 114(2), pp. 577–99.

Borge, L.-E. and Rattsø, J. (2002). 'Spending growth with vertical fiscal imbalance: Norway 1880–1990', *Economics and Politics*, vol. 14(3), pp. 351–73.

Bound, J., Griliches, Z. and Hall, B. H. (1986). 'Wages, schooling and IQ of brothers and sisters: do the family factors differ?', *International Economic Review*, vol. 27(1), pp. 77–105.

Bureau of the Census (1994). *Geographic Areas Reference Manual*, Washington, D.C.: U.S. Dept. of Commerce, Economics and Statistics Administration, Bureau of the Census.

Byfuglien, J. and Langen, O. R. (1983). 'Grunnkretser, tettsteder og menigheter: Dokumentasjon 1980', Rapport 13/83, Statistisk Sentralbyrå.

Conley, T. G. and Topa, G. (2002). 'Socio-economic distance and spatial patterns in unemployment', *Journal of Applied Econometrics*, vol. 17(4), pp. 303–27.

Dearden, L., Machin, S. and Reed, H. (1997). 'Intergenerational mobility in Britain', *Economic Journal*, vol. 107(440), pp. 47–66.

Durlauf, S. N. (2001). 'The memberships theory of poverty: The role of group affiliations in determining socioeconomic outcomes', in (S. H. Danziger and R. H. Haveman, eds.), *Understanding Poverty*, chap. 11, pp. 392–416, Cambridge, MA: Harvard University Press.

Edin, P.-A., Fredriksson, P. and Åslund, O. (2003). 'Ethnic enclaves and the economic success of immigrants – evidence from a natural experiment', *Quarterly Journal of Economics*, vol. 118(1), pp. 329–57.

Ermisch, J. and Francesconi, M. (2001). 'Family matters: impacts of family background on educational attainments', *Economica*, vol. 68(2), pp. 137–56.

Falch, T. and Tovmo, P. (2003). 'Norwegian local public finance in the 1930s', *European Review of Economic History*, vol. 7, pp. 127–54.

Gibbons, S. (2002). 'Neighbourhood effects on educational achievement: evidence from the Census and National Child Development Study', mimeo, Centre for Economics and Education.

Gibbons, S. (2003). 'Paying for good neighbours: estimating the value of an educated community', *Urban Studies*, vol. 40(4), pp. 809–33.

Gibbons, S. and Machin, S. (2003). 'Valuing English primary schools', *Journal of Urban Economics*, vol. 53(2), pp. 197–219.

Ginther, D., Haveman, R. and Wolfe, B. (2000). 'Neighborhood attributes as determinants of children's outcomes: how robust are the relationships?', *Journal of Human Resources*, vol. 35(4), pp. 603–42.

Griliches, Z. (1979). 'Sibling models and data in economics: beginnings of a survey', *Journal of Political Economy*, vol. 87(5), pp. S37–64.

Hægeland, T., Klette, T. J. and Salvanes, K. G. (1999). 'Declining returns to education in Norway? Comparing estimates across cohorts, sectors and time', *Scandinavian Journal of Economics*, vol. 101(4), pp. 555–76.

Haider, S. and Solon, G. (2004). 'Life-cycle variation in the association between current and lifetime earnings', mimeo, where?.

Jargowsky, P. A. (1996). 'Take the money and run: economic segregation in US metropolitan areas', *American Sociological Review*, vol. 61(6), pp. 984–98.

Katz, L. F., Kling, J. R. and Liebman, J. B. (2001). 'Moving to opportunity in Boston: Early results of a randomized mobility experiment', *Quarterly Journal of Economics*, vol. 116(2), pp. 607–54.

Kremer, M. (1997). 'How much does sorting increase inequality?', *Quarterly Journal of Economics*, vol. 112(1), pp. 115–39.

Kunz, J., Page, M. E. and Solon, G. (2003). 'Are point-in-time measures of neighborhood characteristics useful proxies for children's long-run neighborhood environment?', *Economics Letters*, vol. 79(2), pp. 231–37.

Langen, O. R. (1975). 'Folketellingskretsene i Norge 1960 og 1970: Sammenliknbare kretsenheter', mimeo Geografisk institutt, Norges Handelshøyskole og Universitetet i Bergen, geografiske studier av utkantstrøk i Norge, Arbeidsrapport nr. 6.

Møen, J., Salvanes, K. G. and Sørensen, E. Ø. (2003). 'Documentation of the linked employer-employee dataset at Norwegian School of Economics', mimeo, Norwegian School of Economics.

Oreopoulos, P. (2003). 'The long-run consequences of living in a poor neighborhood', *Quarterly Journal of Economics*, vol. 118(4), pp. 533–75.

Page, M. E. and Solon, G. (2003). 'Correlations between brothers and neighboring boys in their adult earnings: the importance of being urban', *Journal of Labor Economics*, vol. 21(4), pp. 831–55.

Solon, G. (1999). 'Intergenerational mobility in the labor market', in (O. Ashenfelter and D. Card, eds.), *Handbook of Labor Economics*, vol. 3a, chap. 29, pp. 1761–800, Amsterdam: North-Holland.

Solon, G., Corcoran, M., Gordon, R. and Laren, D. (1991). 'A longitudinal analysis of sibling correlations in economic status', *Journal of Human Resources*, vol. 26(3), pp. 509–34.

Solon, G., Page, M. E. and Duncan, G. J. (2000). 'Correlation between neighboring children in their subsequent educational attainment', *Review of Economics and Statistics*, vol. 82(3), pp. 383–92.

Statistics Norway (2001). *Statistical Yearbook 2001*, Statistics Norway.

Topa, G. (2001). 'Social interactions, local spillovers and unemployment', *Review of Economic Studies*, vol. 68(2), pp. 261–95.

Vassenden, E. (1995). 'The register of the level of education: a product of the integration of other registers, census and surveys', Statistisk sentralbyrå, Notater 95/4.

Vassenden, K. (1987). 'Folke- og boligtellingene 1960, 1970 og 1980: Dokumentasjon av de sammenlignbare filene', Statistisk sentralbyrå, Notater 87/2.

SWEDISH ECONOMIC POLICY REVIEW 10 (2003) 143-169

The impact of a primary school reform on educational stratification: A Norwegian study of neighbour and school mate correlations

Oddbjørn Raaum, Kjell G. Salvanes and Erik Ø. Sørensen*

Summary

■ School quality is hard to define and measure. It is influenced by not only school expenditures, but also characteristics that are hard to measure like norms and peer effects among teachers and pupils. Furthermore, family background and community characteristics are important in explaining educational outcomes. In this paper we study the composite effect of primary schools and neighbourhoods on adult educational attainment controlling for family characteristics. Instead of identifying the effect of specific neighbourhood and school characteristics on educational attainment, we focus on correlations in final years of schooling among neighbouring children and school mates. We find a clear trend of declining influence of childhood location over the 24 year period (birth cohorts 1947-1970). Then we ask whether a change in the compulsory school law extending the mandatory years of education, can explain this pattern. We find some effect of the primary school reform on the change in the neighbourhood effect. Motivated by the fact that neighbouring children typically go to the same school, we estimate school mate correlations for children born in the 1960s. The overall impact of factors shared by children who graduated from the same school at the age of 15/16 is negligible. The variation in "school quality" and the impact of peers on final educational attainment seem to have been very limited in Norway.■
JEL classification: I21, J13, R23.
Keywords: Families, neighbours, schools, educational reforms, peer-effects.

* *Oddbjørn Raaum is Senior Researcher at the Ragnar Frisch Centre for Economic Research. Kjell G. Salvanes is Professor at the Department of Economics, The Norwegian School of Economics and Business Administration. Erik Ø. Sørensen is Ph.D. student at Department of Economics, The Norwegian School of Economics and Business Administration.*

SWEDISH ECONOMIC POLICY REVIEW 10 (2003) 143-169

The impact of a primary school reform on educational stratification: A Norwegian study of neighbour and school mate correlations

Oddbjørn Raaum, Kjell G. Salvanes
and Erik Ø. Sørensen*

There is a controversy both among researchers and among the public on whether school quality matters, how (much) it matters, why it matters, and for which outcomes it matters. Empirical results from many countries have shown that school resources only have a modest impact on student achievement, but a relatively stronger impact on adult educational attainment.[1] School quality is hard to define and measure. It is not only influenced by school expenditures, but also characteristics that are hard to measure like norms, attitudes and peer effects among teachers and pupils (Hoxby, 2000). Furthermore, family background is important, and it has become clear that community characteristics—peer effects and neighbourhood institutions—are important in explaining educational attainment and adult earnings (Solon, et al. 2000; Page and Solon, 2003; Raaum, Salvanes and Sørensen, 2001).

However, few studies focus on the fact the primary/lower secondary school constitutes an important factor shared by children growing up in the same neighbourhood. The three factors, family, neighbourhoods (both as peer influence and institutions), and schools, probably also interact strongly as inputs in the human capital production function. Therefore, the causal effects of the three factors

* *We thank Peter Fredriksson, a referee and the editor for valuable comments and suggestions.*

[1] School inputs measured by e.g. expenditures and teacher-pupil ratios appear to have modest, if any, effect on student achievement such as marks and test scores (see Hanushek, 2003, and Krueger, 2003, for US results, and Bonesrønning, 2003, for Norway). However, the same type of school resources seem to have a stronger impact on post-school outcomes like final educational attainment and earnings, although these findings are controversial (Betts, 1996; Dearden, Ferri and Meghir, 2002; Dustmann, Rajan and Van Soest, 2003).

are hard to disentangle, partly because of family sorting into neighbourhoods and schools. In the present paper, we try to analyse the composite effect of primary schools and neighbourhoods, and attempt to assess their impact on the variance in adult educational attainment. Instead of identifying the effect of specific neighbourhood and school characteristics on educational attainment, we focus on correlations in the final years of schooling among neighbouring children and school mates.

The starting point for our analysis is twofold. First, in a previous study we found that the importance of family background was stable while the effect of neighbourhoods on educational attainment and earnings is significantly lower for the 1955-65 birth cohort as compared to individuals born 1945-55 (Raaum, Salvanes and Sørensen, 2001). Second, in the 1960s—which is the childhood period for which the neighbourhood effect was found to be weakened—a primary school reform took place in Norway, extending the mandatory level of schooling from seven to nine years. Pre-reform, the Norwegian school system required children to attend school from the age of seven to the age of fourteen. After the reform, this was extended to the age of sixteen by adding two more years of mandatory education. The reform took place over a twelve-year period with different municipalities adopting the new school system at different times, allowing for time variation as well as regional variation. Evidence in Aakvik, Salvanes and Vaage (2003) suggests that this reform increased the participation in the above mandatory education as well as the returns to education.[2] They also found that the importance of family income was slightly weakened for post-reform students as compared to pre-reform students.

Our approach in this paper is to use a unique data set for Norway on neighbourhoods, schools and parental background to analyse whether the school reform also had an impact on equalizing the opportunity across neighbourhoods.[3] In addition to aims such as increasing the minimum level of education, and smoothing the transition to higher education, an important aim was also to increase equal-

[2] See Meghir and Palme (2003) for an analysis of the similar Swedish reform that took place in the 1950s. They also find that the reform had an impact on participation rates in higher education as well as reducing the impact of family background.
[3] See Oreopoulos (2003), Lochner and Morietti (2001) and Pischke (2003) for other examples of analysing social returns, as opposed to private returns of educational reforms.

ity of opportunity along socio-economic and geographic dimensions. It is this latter aspect that we analyse in this paper. The question is whether the school reform reduced the importance of the local neighbourhood. This type of primary school reform took place at about the same time in many other European countries and we think that Norway is a good case for analysing social returns of primary educational reforms, since the potential impact is expected to be stronger and thus easier to measure in the case of Norway. It has been pointed out that the Norwegian reform along with the Swedish reform went further both in the unification of the comprehensive school system as well as in promoting equality of opportunities (Leschinsky and Mayer, 1990). We then analyse the effect of schools as a part of neighbourhood effects by estimating school mate correlations over time, both as unadjusted correlations and controlled for family sorting. A school mate correlation is an overall measure of neighbourhood and different types of school effects, including school resource and peer effect. Again the question is whether school mate correlations have been reduced over time.

The rest of the paper is organized as follows. In Section 1, we describe our approach. In Section 2 the data set, variable definitions and the educational reform are described. Section 3 provides the empirical results and the last section gives some concluding remarks.

1. Neighbour and school mate correlations

In order to study the impact of schools on adult educational attainment as measured by years of education, and to disentangle the family effects and neighbourhood effects, we use a variance decomposition approach. The idea is simple. If childhood neighbourhood have long-lasting effects on welfare, a resemblance in adult outcomes will appear among persons who grew up in the same local community. The same line of reasoning applies to schools and children who graduated from the same institution. Our empirical neighbour (school mate) correlation is an estimate of the proportion of the variance in years of schooling explained by factors shared by neighbouring children (school mates).

In order to illustrate the variance decomposition approach, we use a simple framework suggested by Solon et al. (2000). Let y_{cfi} be the years of education, for sibling i in family f in neighbourhood c.

$$y_{cfi} = \beta' Z_c + \alpha' X_{fc} + \varepsilon_{cfi}, \tag{1}$$

where X_{fc} is a vector of all family characteristics that influence years of education, Z_c contains all the neighbourhood characteristics, and ε_{fi} represents unrelated individual factors orthogonal to both family and neighbourhood effects.[4] The total variance in years of schooling can be decomposed as:

$$\begin{aligned} \operatorname{var}(y_{fi}) = {} & \operatorname{var}(\beta' Z_c) + \operatorname{var}(\alpha' X_{fc}) + 2\operatorname{cov}(\beta' Z_c, \alpha' X_{fc}) \\ & + \operatorname{var}(\varepsilon_{fi}). \end{aligned} \tag{2}$$

We are looking for the relative influence of neighbourhoods on schooling, i.e. $\operatorname{var}(\beta' Z_c)/\operatorname{var}(y_{fi})$. Empirically, we use the observed co-variance in educational attainment among neighbouring children from different families. This covariance, using (1), is given by

$$\begin{aligned} \operatorname{cov}(y_{fi}, y_{f'i'}) = {} & \operatorname{var}(\beta' Z_c) + \operatorname{cov}(\alpha' X_{fc}, \alpha' X_{f'c}) \\ & + 2\operatorname{cov}(\beta' Z_c, \alpha' X_{fc}). \end{aligned} \tag{3}$$

As illustrated in (3), the neighbour covariance contains more than the variance of neighbourhood effects. The second term represents clustering of similar families in neighbourhoods. As families typically sort themselves into neighbourhoods, resemblance in outcomes of children growing up in the same local community (or school) will also reflect similar family backgrounds. The third term reflects the extent to which families are non-randomly distributed across neighbourhoods. We expect that advantaged families sort into advantaged neighbourhoods, reinforcing the impact of a non-random distribution of families on observed neighbour correlation. In the case of school mate correlations, compensating resource allocation across schools will tend to reduce—and possibly even reverse—the positive association between family and school effects.

[4] Since X and Z are latent vectors that include all relevant variables, the residual is orthogonal to both.

148

Empirically, we can estimate the part of $\alpha' X_{fc}$ related to *observed* family characteristics, and adjust the correlation, $\mathrm{cov}(y_{cfi}, y_{cf'i'})/\mathrm{var}(y_{cfi})$, by subtracting the covariance in predicted family effects (divided by the variance). However, since we control only for observed family characteristics, our estimated neighbour correlation represents an upper bound on the neighbourhood effects (see Altonji, 1988; Solon et al., 2000; and Page and Solon, 2000).[5] Obviously, the correlation in adult outcomes among persons who spent their childhood in the same local community cannot tell why neighbourhoods matter. It includes the joint effects of the distribution of characteristics (Z's) and their causal effects (β's).

"Neighbourhood effects" is a label for a variety of different mechanisms. The attitudes and behaviour of peers, the existence and enforcement of social norms as well as local institutions vary across neighbourhoods. Our focus is on the role of the primary school[6] as a potentially important factor shared by neighbouring children. Disentangling the impact of schools from other neighbourhood characteristics is hard as we do not have any reliable information (or assumptions) on the sorting of neighbours across schools, e.g. why neighbouring children go to different schools.

Our approach is less ambitious. First, we estimate the trend in neighbour correlations over a 25-year period, i.e. birth cohorts 1947-1970, with and without family background adjustment. As a by-product, we report estimates of the trend in intergenerational educational mobility. Second, we focus on specific birth cohorts, 1947-1956, that were affected by the primary school reform during the 1960s. We exploit this by estimating neighbour correlations in adult educational attainment by birth cohort for individuals as a function whether they lived in reform or non-reform municipalities. The idea is to assess whether the declining impact of neighbourhoods on educational attainment can be attributed to the introduction of the new school system. Finally, we look at school effects by means of school mate correlations. Resemblance in educational attainment among

[5] Variance decomposition to obtain the upper bound of effect of observed and unobserved effects may be preferred to regression analysis where studies often report unstable and small effects of community characteristics when these are directly included in the estimation equations of adult earnings or educational attainment (for an overview, see Ginther, Haveman and Wolfe, 2000).

[6] By primary school, we mean institutions responsible for compulsory schooling. It includes what is frequently called (lower) secondary levels.

children graduating from the same primary school will reflect the total contribution of school characteristics, including resources and composition of pupils. As similar families tend to cluster in schools, parental background adjustment is needed to tighten the upper bound on school effects. On the other hand, it is not so obvious that disadvantaged families sort into disadvantaged schools, as the allocation of school resources tends to favour schools with children in need of special treatment. Unfortunately, data on primary school attendance are not available for children born before 1960.

2. Data and school institutions

2.1. Families, neighbourhoods and school mates

The data set has been put together from sources provided by Statistics Norway (Møen, Salvanes, and Sørensen, 2003). The data include linked administrative data covering most of the Norwegian residents. We also have national censuses for 1960 and 1970 (Vassenden, 1987). Using a unique personal identifier given to all Norwegian residents by the national population register, we can link records from these data sets. We use a set of household and census tract identifiers in the census to identify families and place of residence during childhood. For the 1959-1970 birth cohorts, we have added which primary school they graduated from. The censuses also contain family background variables such as parents' education. The administrative register contains information on adult taxable income (excluding capital gains) and educational attainment. The linking of administrative to census data is not perfect, but for the subset of individuals we consider in this paper, more than 90 per cent can be linked across these datasets for the older cohorts, while the degree of linking is close to 100 per cent for younger non-immigrants. The main reason for non-linking is that the central register of residents based its first records on the census of 1960, and among those who left home before 1960, little was done to refine the information on parents. We have to drop some additional individuals with incomplete information on residence. Vassenden (1987) documents the construction and linking of the census files, while Statistics Norway (2001) documents the central register of education.

Neighbourhood is defined as census tract in 1960 or 1970. The total number of tracts increased from 7996 in 1960 to 8818 in 1970, with most of the increase in urban areas. The average tract populations were 464 and 439 respectively, and these tracts were considerably smaller than those of most other country censuses of the time (Langen, 1975). With the single-year cohorts we use, the average number of individuals ("neighbours") per neighbourhood on which we have information varies from 6.1 to 8.4, with median neighbourhoods of 4 and 5 individuals.

School mates are individuals who graduated from the same school when leaving compulsory education (age 15/16 typically). The schools are larger than the neighbourhoods, with average cohorts of 62 students (median 49) in the 1959 cohort, with a trend toward smaller schools; in the latest cohort for which we have a full year, 1969, the mean graduating class has 55 students (median 41).[7]

We observe the neighbourhood in which children live at one point in time. Because families move, the neighbourhood at a single point in time may not accurately represent the environment children grew up in. On the other hand, people may move between very similar neighbourhoods. In a previous paper (Raaum, Salvanes and Sørensen, 2001), we examined the differential outcomes among those who stayed and those who changed location between the 1960 and 1970 census (using the list of comparable tract aggregations provided by Langen, 1975). We found that with respect to neighbour-correlations in adult educational achievement, this factor does not seem to cause major biases.

There are 451 municipalities in the 1970 census, and most of these have at least one school each, only a few have joint schools with neighbouring municipalities. In 1974, 247 municipalities had only one school, but there are 827 schools in all, which gives an overall average of 1.96 schools per municipality. Typically, a school district contains a number of census tracts and, by regulation, a census tract should not cross school district boundaries although this policy was more strictly enforced in rural than in urban areas (Byfuglien and Langen, 1983). Since some time passed between the census of 1970 and our observations of graduations, which appear from 1974 and onwards, internal migration makes it difficult to examine the map from census tracts to school districts in great detail. Noise induced by migration is correlated with the size of the school district, but the median school dis-

[7] The 1970 cohort is truncated since we have no information on people born after the date of the census (November 1, 1970).

trict had, as of 1974, graduates from 15 census tracts, whereas the 25th percentile school district had graduates from 11 tracts and the 75th percentile had 23 tracts represented.

Our measure of adult educational attainment in our main sample is taken from the register of the level of education maintained by Statistics Norway (Statistics Norway, 2001). This register provides a detailed code of the highest completed education, the completion date and to how many years of schooling the highest completed education corresponds. For individuals with no recent education, their level of education as of the 1970 census is recorded.

Information on the educational attainment of parents is different. The 1960 census data on parents contain only categorical coding of types of education. We have transformed the categorical education codes into years of education, using a two-step procedure. A first step maps 1960 census codes to 1970 census codes, using repeated observations of the same individuals in the two censuses. A second step maps 1970 codes into years of schooling, using the oldest observations in the central register of education. See Raaum, Salvanes and Sørensen (2001) for the details of this procedure.

2.2. The Norwegian mandatory school reform in the 1960s

In 1959, the Norwegian Parliament passed a law on mandatory schooling and the new compulsory 9 years of schooling were gradually implemented across the country over the years 1960 to 1972. This school reform extended the number of compulsory years of schooling from 7 to 9, keeping the school-starting age constant at 7. It also unified the education system beyond the age of 15/16. Before the reform, two years of junior high school preparing for senior high school were possible to obtain in some municipalities, but pupils in other areas had to move to another municipality to attend post-compulsory schools. The nine years in the new system were divided into two levels; first six years of primary school, then three years of lower secondary school which prepared for high school. Hence, for more than a decade, the Norwegian compulsory school was divided into two separate systems. The first cohort that was involved in the reform was the one born in 1947 and the last cohort that went through the old system was born in 1959.

The aims of the reform, explicitly stated in several governmental papers, were to increase the minimum level of educational attainment by extending the number of years of compulsory education, to smooth the transition to higher education, and finally to enhance equality of opportunities, both along the socio-economic and the geographical dimension.

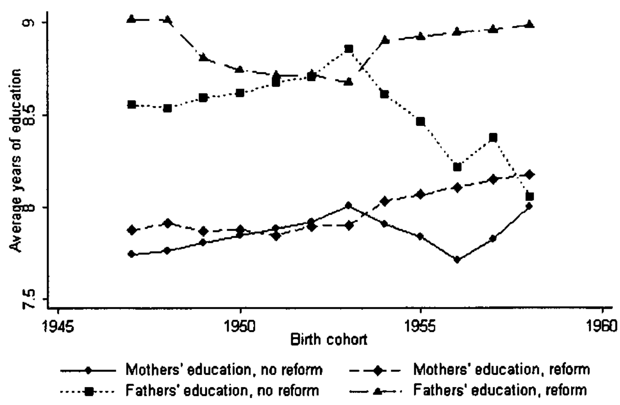
Implementation process of the reform

Under the law of 1959 for mandatory schooling, each municipality was invited to apply to a committee under the Ministry of Education to implement the reformed school system for the whole municipality. This application should include a plan for the new school in terms of buildings and funding, although the extra costs of teachers and buildings were provided by the state. The criteria for being selected among the applicants by the committee are not clear. However, the committee wanted to cover different types of communities, making the sample of implementing municipalities representative for the country and also the plans for buildings, teaching resources etc should be acceptable (Telhaug, 1969; Mediås, 2000).

We are assessing changes in neighbourhood effects (a relative measure) and not levels of education. Thus, we are less vulnerable to the problem of whether reform adoption was random in terms of school participation above the mandatory years of education. However, the question that is of course of interest also in our case is whether municipalities that have implemented the new system, at any given time (or for any given birth cohort), do not vary systematically from those who still kept the old school with 7 years of compulsory schooling. When comparing municipalities by reform status, systematic unobserved heterogeneity may bias our results. For instance, did the richest municipalities implement the reform first? Was it the cities? Or did municipals in poor rural areas implement the reform first since there were obvious economic incentives for implementing the reform? In the public debate from the 1950s and 1960s, it was claimed that the old educational system with more streaming, prepared better for high school and university studies than the new system, indicating that the rich and city areas perhaps implemented the reform late. It was also claimed in the public debate at the time that 9 years of mandatory schooling were not necessary in many rural communities, since fishing and farming were the main industries and those did not require 9 years or higher education.

We are not checking these hypotheses carefully in this paper, only presenting some indication of a possible relationship between the average years of parental schooling, by the birth cohort of their child and the reform status displayed in Figure 1. The figure suggests that the unconditional transition (probability) was positively correlated with the educational attainment of the parents. In Aakvik, Salvanes and Vaage (2003), a detailed analysis of the process of allocating the reforms to municipalities is undertaken. As indicated from Figure 1, the case is not completely clear, but a more detailed analysis did not find support for a systematic allocation of the reform to municipalities.

Figure 1. Parental years of education (by primary school reform status and birth cohort)



Identification of reform status

Information on what type of primary school people attended is only available for those who never continued schooling above the mandatory years, so it is necessary to classify the type of primary education based on municipality of residence in the censuses of 1960 and 1970. It is, however, not an easy task to find municipality level information on reform implementation. The most authoritative list is Ness (1971), but this list is organized by municipalities in 1970. A series of municipality mergers and boundary adjustments in 1965-66 make it difficult to fix a point in time for the reform based on 1960 municipality for the later part of the 1960s. We want to concentrate on finding a date of implementation using the 1960 municipalities; since a 1970 municipality can include several 1960 municipalities with different dates of implementation and thus, it is more difficult to fix a unique implementation year for the 1970 municipalities.

We use a classification scheme based on administrative data on adult educational achievement, focusing on those who left school with only primary education, let us call these people the "dropouts". For each 1960 municipality, we follow the cohorts of those who lived there at the time of the 1960 census. For each year, we calculate the share of dropouts from the old system and the share of dropouts from the new system. We use these dropout rates to calculate two candidate measures of reform date: The first when the dropouts from the old system stop appearing, and the second when the dropouts from the new system start showing up.

Since we must allow for some migration, we cannot simply use indicators of whether there are any dropouts at all as measures of school type. Such a scheme would be much too sensitive to internal migration of even a single individual who moved and dropped out in a municipality with a different implementation date than the one he left. This problem would be particularly important for dating the reform in the larger municipalities, since they receive the large number of migrants. In order to get around this, we need to measure the number of dropouts relative to the population of potential dropouts, and we need to set a positive cut-off rate to allow for some measurement error. We also want to avoid that this measurement error is systematically related to the schooling pattern in the municipalities, so we cannot use a uniform cut-off rate across all municipalities. Instead, we calculate municipality-specific "normal rates" of dropout based on the dropout rates of the 1946-1948 cohorts, which were not exposed to the reform. When the dropout rate from the old system falls below 50 per cent of this "normal" rate, we have the first candidate date of the implementation of the reform. Similarly, we calculate such normal dropout rates from the new system using the 1957-1959 cohorts which we know with certainty went through the new system. The year the rate of dropouts from the new system reaches 50 per cent of this second normal rate is our second candidate date of reform implementation.

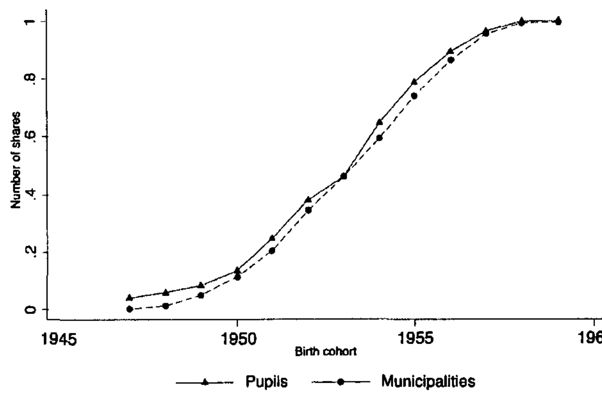
When the two candidate measures agree on what year the reform was implemented, we use this as the year of implementation. Should there be a gap of one or two years between the two candidate measures, such that it would seem that the old system closed before the new one opened, we use the second candidate measure since this is most resistant to a secular decrease in the dropout rate. Should there instead be an overlap of one year between the two candidate measures, such that it seems that the old system and the new system coexisted for a year, we tried to check all larger municipalities (with more than 100 students) against the list in Ness (1971) and local informants. For smaller municipalities with one-year overlaps, we have randomly assigned one of the candidate years. The remaining municipalities, for which none of these methods worked, have been dropped from the sample. While there will certainly be some measurement error in our reform date taken as a flow indicator, we believe that the measurement error in the stock of reformed and non-reformed municipalities for a given year is small.

This method provides a year of implementation for 545 out of 728 municipalities. Table 1 displays the relative importance of the various rules in assigning an implementation date. The slow and gradual implementation of the reform is illustrated in Figure 2. Table A4 in the Appendix presents descriptive statistics for the included and non-included neighbourhoods. As we can see, there is very little difference. In our analysis below, we only consider birth cohorts where the smallest of the reformed and non-reformed group constitute at least 5 per cent of the students, and we therefore exclude the 1946-1947 and the 1957-1959 cohorts.

Table 1. Procedures of reform year identification

	Share of municipalities	Share of pupils
The two indicators coincide	.398	.555
One-year gap	.143	.125
Two-year gap	.059	.042
Manual inspection	.029	.078
Random assignment	.116	.071
Undecided, not used	.255	.129

Figure 2. Accumulated shares of after-reform municipalities and pupils



3. Results

Neighbour and school-mate correlations are estimated using the full list of all unique pairs within neighbourhoods or schools that are not also siblings; see Solon et al. (2000). Correlations are reported separately for each birth year, in order to distinguish between neighbourhoods located in pre- and post-reform municipalities. If we expanded the number of birth cohorts, each neighbourhood would consist of children who went to different school systems. One might argue that children are affected by the attitudes and behaviour among older peers and not only by those of equal age. However, those born in the same year would be exposed to the same environment, e.g. have the same older role models. Detailed results are reported in Appendix Tables A1-A3.

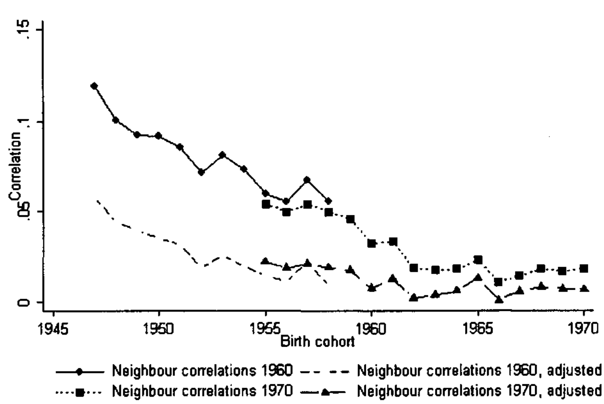
3.1. Trend in the effects of childhood neighbourhood and parental education

Figure 3 displays the correlations in educational attainment among neighbouring children by birth cohort.[8] The neighbourhoods of the

[8] The standard errors are not displayed, but they are very small and vary around .006, see the Appendix.

1947-1958 cohorts are defined by the 1960 census, while the 1970 census defines the neighbourhoods for the 1955-1970 cohorts. The figure also includes the family background adjusted correlations which subtract the covariance component arising from sorting on observed family characteristics (i.e. parental education). The correlations are substantial, around .1, for the cohorts born in the late 1940s and early 1950s. There is a clear trend of declining correlations until around the 1962 cohort, but from then onwards, the correlations are basically constant at a level of about .025. Since the estimates using the two alternative neighbourhood definitions for the "overlapping" birth cohorts, 1955-58, are basically the same, the lower correlations in the 1960 cohorts cannot simply reflect a change in the definition of neighbourhoods.

Figure 3. Neighbour correlations by birth cohort



Apparently, correlations in Figure 3 are small and some may find them negligible. However, if we convert a correlation estimate of .1 into "level effects" in years of schooling, we get a standard deviation of neighbourhood effects which amounts to about .95 years.[9] A correlation of .03 corresponds to a standard deviation of .5 years of schooling. Consequently, even seemingly negligible correlations are non-trivial. For comparison, a correlation of .4, which is the typical

[9] By rearranging (3) and using the observed standard deviation in schooling, which is about 3.

number for Norwegian siblings, corresponds to a standard deviation of effects of 1.9 years of schooling.

Figure 3 also reveals that family sorting matters. In order to adjust for parental education, we regress educational attainment on schooling years of the father and mother and neighbourhood dummies. Subtracting the covariance of predicted family effects from the total covariance and dividing by the total variance of educational attainment, we get the adjusted neighbour correlations. When correlations are adjusted for parental education, the estimates are reduced by more than fifty percent. While the neighbour correlations for the cohorts in the late 1940s and early 1950s remain significant, at around .04, they drop steadily over time and are close to zero from the 1960-cohorts onwards. As even the family adjusted correlations can be seen as upper bounds on the neighbourhood effects, we conclude that the impact of childhood community on adult educational attainment is negligible for Norwegians who are today in their thirties and early forties.

The declining neighbourhood effects may reflect that sorting on unobserved family characteristics has become less severe over time. This explanation can be checked by looking at how adult education is distributed within and between neighbourhoods over time, since we expect the sorting on the basis of parental education to be the same as that on unobserved characteristics. Table 2 is taken from Raaum, Salvanes and Sørensen (2001) and shows that the between-neighbourhood component has become more important over time, indicating that sorting has been more, rather than less, severe.

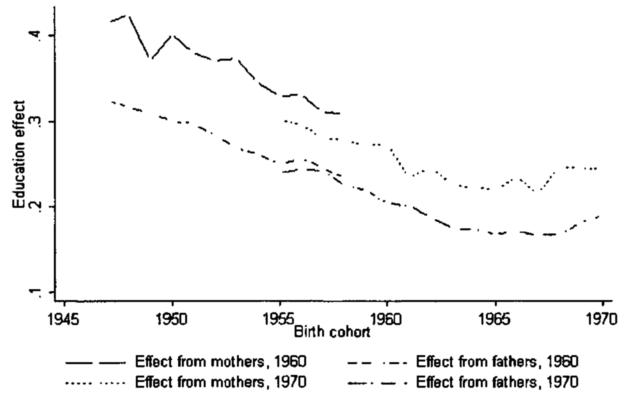
Table 2. Degree of neighbourhood sorting

	Mother's education		Father's education	
	1945-55	1955-65	1945-55	1955-65
mean	8.005	8.678	8.771	9.503
$\hat{\sigma}_u$	.611	.846	.780	1.314
$\hat{\sigma}_e$	1.578	1.814	1.873	2.505
$\hat{\rho} = \hat{\sigma}_u^2 / \left(\hat{\sigma}_u^2 + \hat{\sigma}_\varepsilon^2\right)$	.130	.179	.171	.216

Note: The decomposition of the variance of parental schooling. Estimates from the fixed effect regression $E_{it} = \overline{E} + u_t + \varepsilon_{it}$ (neighbourhood fixed effects). Sample is restricted to parents aged 30-50 at the time of the censuses. This table is taken from Raaum, Salvanes and Sørensen (2001).

The family adjustment is based on cohort-specific estimates of the relation between the schooling years of parents' and children. Figure 4 displays the estimated regression coefficient of the schooling years of the mother and father. An interaction term turns out negative and the coefficients are evaluated at the mean for fathers and mothers. Two striking results appear. First, there is a clear trend of declining relation between educational attainment of parents and child, suggesting that intergenerational educational mobility has increased, in accordance with Bratberg, Nilsen and Vaage (2002). Second, the "effect" of the mother's education seems to be the stronger.

Figure 4. Effects of parental education on children's schooling (by birth cohort)



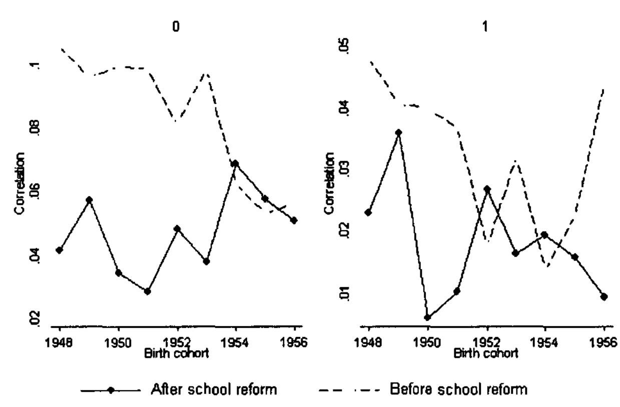
3.2. Neighbour correlations by primary school reform status

For each of the 1948-1956 birth cohorts, we classify individuals as "before- or after-reform" according to the reform-status of the municipality in which their neighbourhood is located. Neighbour correlations are then estimated separately by cohort and reform status. This exercise is motivated by the pattern of declining neighbour correlations; if the primary school reform lowered the impact of childhood location, we expect to find a lower correlation among neighbouring children who went to the new school system. Consequently, as more children were entering the new school, the overall neighbour correlation would drop as a result of the reform.

The neighbour correlations are displayed in the left panel in Figure 5, while the family adjusted estimates are shown in the right panel. First, looking at the left panel, we see that the after-reform correlations are all lower than the before-reform correlations during the first seven years (incl. the 1953 cohort). By 1953, about 50 per cent of the cohort lived in municipalities which had implemented the new school system. Thereafter, the correlations of two groups are basically the same. We also see that the trend of declining correlations, with the exception of the 1953-cohort, remains when we consider the before-reform neighbourhoods. No such trend is found for the after-reform individuals.

Second, the right panel shows that the difference according to reform status drops significantly when we adjust for parental education. Although the estimated neighbour correlations are higher in the before-reform municipalities in seven of nine cases, there is no clear pattern. There is a tendency to lower post-reform correlations in municipalities with an early implementation. This is restricted to the 1947-1951 cohorts and the fraction of pupils in the new school is less than 25 per cent in these cohorts. At most, the primary school reform implemented throughout the 1960s had a modest impact on the overall trend of declining neighbourhood effects.

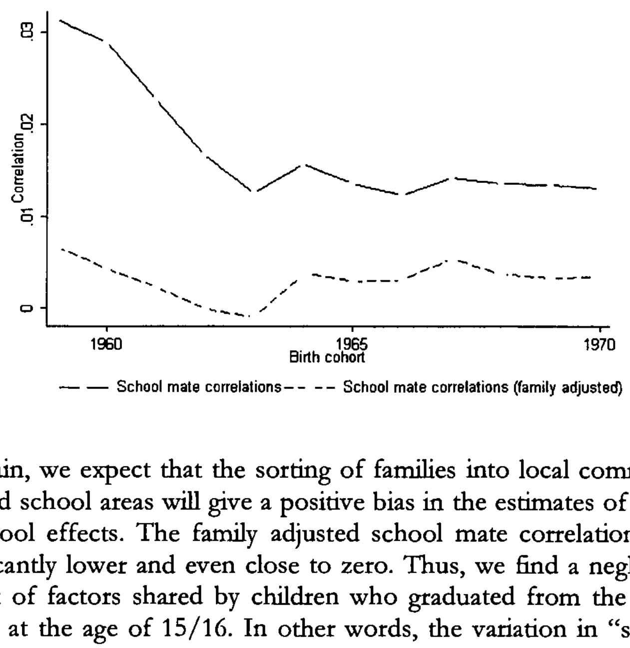
Figure 5. Neighbour correlations (by primary school reform status and birth cohort)



3.3. School mate correlations

A lower secondary school identifier is available from 1974 onwards, enabling us to construct school mates defined as children born between 1959 and 1970 who graduated from the same school around the age of 15/16. All went to the new system with nine years of compulsory schooling. Figure 6 displays correlations in years of schooling among school mates, by birth cohort. The upper line shows the unadjusted correlations and we recognize the pattern of declining correlations found among neighbouring children in the pre-1962 cohorts. We note, however, that the school mate correlations are significantly lower than the corresponding neighbour correlations.

Figure 6. School mate correlations (by birth cohort)



Again, we expect that the sorting of families into local communities and school areas will give a positive bias in the estimates of overall school effects. The family adjusted school mate correlations are significantly lower and even close to zero. Thus, we find a negligible impact of factors shared by children who graduated from the same school at the age of 15/16. In other words, the variation in "school quality" and the magnitude of peer-effects seem to be very small. This is consistent with the negligible neighbour correlations of the same cohorts, and also the low levels of "between-school" variance typically found in studies of student performance distributions (Coleman, 1966; OECD, 2003). One caveat needs to be emphasized. The interpretation of a family adjusted school mate correlation as an upper bound on the school effects is based on the assumption that children of "advantaged" families go to "good schools", i.e. $\mathrm{cov}(\beta' Z_c, \alpha' X_{fc}) \geq 0$. Since school resources are partly distributed in a compensating way, which provides extra resources to schools teaching pupils with specific needs, this assumption may not hold. On the other hand, our family background adjustment is unlikely to account for the total impact of clustering of similar families in schools.

4. Conclusions

This paper has studied the composite effect of primary schools and neighbourhoods on adult educational attainment in Norway, with a particular emphasis on changes over time. We focus on correlations in the final years of schooling among neighbouring children as well as school mates. These correlations measure the proportion of the variance in years of schooling, explained by factors shared by individuals who grew up in the same local community or graduated from the same school at the age 15/16. We do not identify the effects of *specific* neighbourhood and school characteristics, but the correlations measure the relative importance of childhood neighbourhood and school. As such, the measures are closely linked to "inequality of opportunity", where the location of your parents' home affects your adult outcome.

The impact of neighbourhoods on educational attainment has diminished, in accordance with Raaum, Salvanes and Sørensen (2001). Estimating neighbourhood effects for all birth cohorts from the late 1947 to 1970, we find a clear trend of declining correlations until around the 1962 cohort. From then onwards, the correlations are basically constant and close to zero when we adjust for family sorting into local communities.

We single out the primary school reform gradually introduced during the 1960s as a potential explanation, because primary schools constitute an important part of the neighbourhoods. The reform extended compulsory schooling from 7 to 9 years, provided a common curriculum for all schools and was aimed at equalizing opportunities across socio-economic and geographical backgrounds. For each of the 1947-1956 birth cohorts, we classify individuals as "before- or after-reform", according to the reform-status of the neighbourhood. The estimated neighbour correlations tend to be higher in the before-reform municipalities, but the difference is reduced when we adjust for parental education. The primary school reform implemented throughout the 1960s cannot fully explain the trend of declining neighbourhood effects in Norway.

Finally, we estimate school mate correlations for children born between 1959 and 1970, looking for the impact of factors shared by children who graduated from the same school at the age 15/16. Effects of school resources and organizational practices, peer effects within schools and local communities are included in this measure. Accounting for family sorting, the school mate correlations are close to zero. Thus, the variation in "school quality" and the impact of peers on final educational attainment seem to have been very limited in Norway, consistent with the negligible neighbour correlations found for the same cohorts.

Focusing on Norwegians presently in their thirties and early forties, their childhood neighbourhood and primary school have had a negligible impact on their educational attainment. Since significant neighbourhood effects are found for those who are ten years older, it seems likely that policy changes have been effective in levelling the playing field across local communities. Even if the effects of the primary school reform are found to be limited, we believe that redistributive policies equalizing spending across municipalities and other educational reforms are likely explanations.

Family background, however, remains an important determinant of educational attainment. The evidence on how family effects have changed over time is mixed. Apparently, the declining relation between educational attainment of children and parents, as well the drop in neighbourhood effects, are both at odds with the stable sibling correlations found in Raaum, Salvanes and Sørensen (2001). As neighbourhood and parental education represent factors typically shared by siblings, we would expect sibling correlations to fall as well. However, alternative measures of intergenerational mobility do not necessarily change in the same direction. Sibling correlations are affected by intra-family resemblance as well as inter-family differences. Imagine that educational reforms induce all "talented" children from "disadvantaged" families (where "talent" is shared by siblings), to continue school and enter higher education. If parental resources only allowed one of the children to enter university in the earlier cohorts, the reforms would reduce intra-family differences which would contribute to a higher resemblance in educational attainment among siblings. This example illustrates the possibility that intra-family resemblance is strengthened, while differences between families are reduced.

In the Nordic countries, access to rich administrative and census data has opened up during the last five to ten years. Matched data on individuals, families, schools and neighbourhoods facilitate new approaches in future studies trying to disentangle the effects of these factors. Good data help considerably, but the real challenge is to establish a framework which enables us to identify behaviour as well as responses to policy changes.

References

Aakvik, A., Salvanes, K.G. and Vaage, K. (2003), Measuring the heterogeneity in the returns to education in Norway using educational reforms, CEPR 3995, Centre for Economic and Policy Research.

Altonji, J.G. (1988), The effects of family background and school characteristics on education and labor market outcomes of relatives, mimeo, Northwestern University.

Betts, J.R. (1996), Do school resources matter only for older workers?, Review of Economics and Statistics 78, 638-652.

Bonesrønning, H. (2003), Class size effects on student achievement: Patterns and explanations, Southern Economic Journal, forthcoming.

Bratberg, E., Øivind A.N. and Vaage, K. (2002), Assessing changes in intergenerational earnings mobility, Working Paper 54, Centre for Economic Studies in Social Insurance, University of Bergen.

Byfuglien, J. and Langen, O.R. (1983), Grunnkretser, tettsteder og menigheter: Dokumentasjon 1980, Rapport 83/13, Statistisk sentralbyrå.

Coleman, J.S. (1966), Equality of Educational Opportunity, US GPO, Washington, D.C.

Dearden, L., Ferri, J. and Meghir, C. (2002), The effect of school quality on educational attainment and wages, Review of Economics and Statistics 84, 1-20.

Dustmann, C., Rajah, N. and Van Soest, A. (2003), Class size, education and wages, Economic Journal 113, F99-F120.

Ginther, D., Haveman, R. and Wolfe, B. (2000), Neighbourhood attributes as determinants of children's outcomes: How robust are the relationships?, Journal of Human Resources 35, 603-642.

Hanushek, E. (2003), The failure of input-based schooling policies, Economic Journal 113, F64-F98.

Hoxby, C. (2000), Peer effects in the classroom, NBER working paper 7867, National Bureau of Economic Research.

Krueger, A.B. (2003), Economic considerations and class size, Economic Journal 113, F34-F63.

Langen, O.R. (1975), Folketellingskretsene i Norge i 1960 og 1970: Sammenliknbare kretsenheter, Arbeidsrapport 6, Geografiske studier av utkanstrøk av Norge, Geografisk institutt, Norwegian School of Economics and Bergen University.

Leschinsky, A. and Mayer, K.A. (eds.) (1990), The Comprehensive School Experiment Revisited: Evidence from Western Europe, P. Lang, Frankfurt am Main.

Lochner, L. and Moretti, E. (2002), The effect of education on crime: Evidence from prison inmates, arrests, and self-reports, NBER Working Paper 8986, National Bureau of Economic Research.

Mediås, O.A. (2000), Fra griffel til PC, Steinkjer, Steinkjer commune.

Meghir, C. and Palme, M. (2003), Ability, parental background and education policy: empirical evidence from a social experiment, Mimeo, Stockholm School of Economics.

Møen, J., Salvanes, K.G. and Sørensen, E.Ø. (2003), Documentation of the linked employer-employee data base at the Norwegian School of Economics, Mimeo, The Norwegian School of Economics and Business Administration.

Ness, E. (red). (1971), Skolens årbok 1971, Johan Grundt Tanum Forlag. Oslo.

OECD (2003). PISA, Literary Skills for the World of Tomorrow, OECD and Unesco, OECD, Paris (http://www1.oecd.org/publications/e-book/960307).

Oreopoulos, P. (2003), The compelling effects of Compulsory schooling: evidence from Canada, Mimeo, Department of Economics, University of Toronto.

Pischke, J.-S. (2003), The impact of length of school years on student performance and earnings: Evidence from the German short school years, CEPR DP 4074, Center for Economic Policy Research.

Raaum, O., Salvanes, K.G. and Sørensen, E.Ø. (2001), The Neighbourhood is not what it used to be, Discussion Paper 36/2001, The Norwegian School of Economics.

Page, M.E. and Solon, G. (2003), Correlations between brothers and neighbouring boys in their adult earnings: The importance of being urban, Journal of Labor Economics, forthcoming.

Solon, G., Page, M.E. and Duncan, G.J. (2000), Correlations between neighbouring children in their subsequent educational attainment, Review of Economics and Statistics 82, 383-392.

Statistics Norway (2001), Utdanningsstatistikk: Dokumentasjon 2000 av den individbaserte utdanningsstatistikken, Norges offisielle statistikk C645, Statistisk sentralbyrå.

Telhaug, A.O. (1969), Den 9-årige skolen og differensieringsproblemet. En oversikt over den historiske utvikling og den aktuelle debatt, Lærerstudentenes Forlag, Oslo.

Vassenden, K. (1987), Folke- og boligtellingene 1960, 1970 og 1980. Dokumentasjon av de sammenliknbare filene. Rapport 87/2, Statistisk sentralbyrå.

Appendix

The correlation estimator is described in Solon et al. (2000). Each neighbourhood or school-mate group consists of many pairs of individuals. Earlier research has indicated that the weighting strategy is not critical, so all these pairs are weighted equally and an ordinary correlation is calculated on this expanded dataset. For calculation of the standard errors, we have used a bootstrap estimator, re-sampling with municipalities as the clustering unit and 300 replications.

Table A1. Trend in neighbour correlations

Birth Cohort	1960 neighbourhoods		1970 neighbourhoods	
	Unadjusted	Family adjusted	Unadjusted	Family adjusted
1947	.1191(.0289)	.0576(.0106)		
1948	.1004(.0179)	.0438(.0056)		
1949	.0924(.0144)	.0392(.0043)		
1950	.0918(.0170)	.0352(.0043)		
1951	.0853(.0153)	.0314(.0041)		
1952	.0713(.0109)	.0187(.0038)		
1953	.0813(.0159)	.0257(.0043)		
1954	.0732(.0128)	.0194(.0032)		
1955	.0596(.0084)	.0140(.0030)	.0534(.0066)	.0220(.0042)
1956	.0554(.0068)	.0112(.0038)	.0498(.0073)	.0187(.0033)
1957	.0673(.0096)	.0215(.0034)	.0542(.0069)	.0213(.0039)
1958	.0552(.0063)	.0089(.0055)	.0495(.0065)	.0189(.0037)
1959			.0460(.0069)	.0173(.0036)
1960			.0321(.0059)	.0074(.0034)
1961			.0332(.0037)	.0122(.0036)
1962			.0185(.0027)	.0018(.0038)
1963			.0175(.0030)	.0036(.0038)
1964			.0177(.0029)	.0058(.0035)
1965			.0229(.0029)	.0128(.0034)
1966			.0106(.0025)	.0008(.0026)
1967			.0141(.0025)	.0055(.0025)
1968			.0177(.0027)	.0080(.0025)
1969			.0166(.0029)	.0070(.0022)
1970			.0181(.0042)	.0066(.0037)

Table A2. Neighbour correlations by reform status

Birth cohort	Post-reform neighbourhoods		Pre-reform neighbourhoods	
	Unadjusted	Family adjusted	Unadjusted	Family adjusted
1948	.0414(.0127)	.0229(.0101)	.1059(.0186)	.0480(.0066)
1949	.0574(.0145)	.0358(.0147)	.0959(.0160)	.0404(.0049)
1950	.0343(.0094)	.0058(.0098)	.0100(.0162)	.0400(.0041)
1951	.0285(.0054)	.0102(.0063)	.0985(.0164)	.0367(.0069)
1952	.0480(.0059)	.0267(.0055)	.0812(.0163)	.0175(.0041)
1953	.0379(.0054)	.0162(.0045)	.0983(.0169)	.0320(.0068)
1954	.0686(.0127)	.0192(.0031)	.0623(.0219)	.0139(.0069)
1955	.0577(.0092)	.0157(.0027)	.0531(.0146)	.0223 (.0077)
1956	.0507(.0072)	.0092(.0036)	.0562(.0101)	.0478(.0089)

Table A3. School mate correlations

Birth cohort	Unadjusted	Family adjusted
1959	.0313 (.0063)	.0067(.0029)
1960	.0289(.0044)	.0043(.0034)
1961	.0226(.0035)	.0023(.0032)
1962	.0167(.0020)	-.0000(.0036)
1963	.0126(.0020)	-.0009(.0038)
1964	.0157(.0015)	.0038(.0031)
1965	.0137(.0016)	.0029(.0029)
1966	.0124(.0015)	.0031(.0022)
1967	.0142(.0018)	.0054(.0024)
1968	.0136(.0017)	.0037(.0027)
1969	.0134(.0019)	.0032(.0033)
1970	.0131(.0018)	.0034(.0025)

Table A4. Descriptive statistics and neighbour correlations for neighbourhoods that were matched and not matched in the pre- and post-reform analysis

Cohort	Average education Matched	Average education No match	Average earning Matched	Average earnings No match	Neighb. correlation Matched	Neighb. correlation No match	Share Matched pupils
1948	11.13	11.14	240133	249901	.04347	.09331	.868
1949	11.23	11.27	248255	256356	.05873	.09616	.865
1950	11.34	11.36	251938	262164	.04561	.09027	.865
1951	11.44	11.47	254454	263525	.06010	.08532	.863
1952	11.45	11.56	255013	270125	.05315	.07689	.862
1953	11.52	11.66	258094	270610	.03496	.07524	.866
1954	11.62	11.68	259919	269817	.04665	.07510	.866
1955	11.67	11.72	264785	271710	.03620	.05701	.865
1956	11.63	11.74	258100	268968	.04451	.05029	.869

Sectoral Choice with Human Capital and Accumulation of Pension Benefits

Erik Ø. Sørensen*
Department of Economics
Norwegian School of Economics and Business Administration

November 2005

Abstract

Universal pension plans and large public-sector workforces affect accumulation and allocation of human capital. The benefit reforms and retraining programs being considered in many countries are likely to affect behaviour in ways that can only be analysed within forward-looking models of lifetime labour supply. Using Norwegian panel data on three birth cohorts, this paper develops and estimate a life-cycle model of public- and private-sector employment. The model of sequential career-choices builds on Keane and Wolpin (1997), extending the accumulation of skills to be sector-specific and allowing unobserved, non-deterministic depreciation of skills. The Norwegian retirement benefits are modelled in a way that builds on the discretisation approach of Rust and Phelan (1997), and identification is aided by exploiting how current career-choices affect future expected benefits. I find important heterogeneity in skill accumulation. The model is used to analyse the effect of a pension reform on sector-specific labour supply. The reform has large effects on labour supply, but the sectoral effects are small.

1 Introduction

This paper constructs a model of life-cycle labour supply in which individuals accumulate sector-specific skills by attending school or gaining work experience, and in which individuals run the risk of losing skills that are not used. This allows

*Department of Economics NHH, Helleveien 30, N-5045 Bergen, erik.sorensen@nhh.no. I am grateful for encouragement and very helpful advice from Christopher Ferrall and Kjell G. Salvanes. I have also received helpful comments from Anne M. Danø.

individuals to endogenously specialise in one of the two sectors, with an expected cost of switching sectors that depends on the skill technology and on the career history. I estimate this model on a panel of three Norwegian birth-cohorts, using the incentives of sector-specific retirement benefits to aid identification. This framework allows evaluation of pension reforms and adult education initiatives. These policies are under consideration in a number of countries as welfare states prepare for ageing populations and react to increasing public pressure to modernise and downsize their public sectors.

Preparing for demographic changes about to come, universal pension plans financed out of current taxation has been the target of reform plans in most developed countries. Such reforms will have consequences for more than government budgets; implicit claims on future benefits constitute a large fraction of households net wealth (Domeij and Klein 2002) and have been claimed to have important effects on life-cycle labour supply (Rust 2001). In addition to the universal plans, large public sector workforces have often accepted generous future supplementary benefits in exchange for a lower level of current earnings. Changes to the universal plans will interact with the public sector supplementary benefits and could have large effects on recruitment to the public sector and labour mobility between the public and the private sector. Quantification of these effects can only be made within a life-cycle model of labour supply.

This paper follows Keane and Wolpin (1997) in modelling the occupational choice problem as a net present value maximisation, with the decision to work being discrete. Individuals take into account the effect current career choices have on future retirement benefits. The paper departs from the baseline model of Keane and Wolpin in making the occupational choice one between the public and private sector rather than a choice between blue-collar/white-collar and military employment. This is motivated by the attempt to analyse policies directed at the public sector employment and benefit reforms, but also by the fact that the public workforce accounts for a large proportion of total employment in most European welfare states. In Norway it accounts for as much as 33% of employment in 2000 (Statistics Norway 2001, Table 244). There is also evidence that wage-education profiles are different in the public and private sector in Norway even after controlling for a large number of other factors (Hægeland, Klette and Salvanes 1999).

The paper shares the goal of Rust and Phelan (1997) of modelling the effect of retirement benefits on behaviour. It also borrows some of the methodology of modelling the dynamics as controlled Markov processes on rather course grids. The state-space can be factored such that some of these Markov processes can be pre-estimated. It differs from Rust and Phelan (1997) in having an explicit model of how skills accumulate and in including career-choices over the full life-cycle: formal schooling and National Service, accumulation of work-experience and retirement at old age.[1] A benefit of this is that a wider range of policies can be studied.[2] The model is to integrate out the effect of past choices on the distribution of unobservables, such that it is not necessary to make the restrictive assumption that all non-transient state-variables can be observed by the econometrician.

The three birth-cohorts of Norwegian males I use span the life-cycle, and following Lee (2005) I use the stochastic structure of the model to integrate out over the unobservables for time that pass before the cohorts under study reaches the window of observation (1986–2000). Observers have claimed that the incentives introduced by the public sector supplementary benefit schemes have the effect of locking people into the public sector when they reach middle age. This difference in retirement benefits between the two sectors provides exogenous variation in incentives that helps identify the model.

The pension system in Norway has two main parts. There is an universal scheme, the "National Insurance" (*Folketrygden*, somewhat resembling the US Social Security), and there are supplimentary systems provided by contracting between employers and employees. The National Insurance is a strongly redistributive program, financed by pay-as-you-go taxation. Even those who have never worked, and therefore not contributed to the pension fund, are entitled to a basic pension and a special supplement guaranteeing a yearly minimum income of about USD 11,000 with little taxation applied. Even above this, the accumulation of claims to future benefits is highly non-actuarial. Employers provide additional supplementary pension schemes. Only the private sector supplement-

[1] Former studies of the Norwegian pension benefits schemes have focused either on forecasting effects of changes in the fine structure of early retirement schemes, and have not modelled the full life-cycle. See for instance Hernæs, Sollie and Strøm (2000); Bratberg, Holmås and Thøgersen (2001); Brinch, Hernaes and Strøm (2001); Jia (2005).

[2] The present paper concentrates on reform of the retirement benefits but future work will look at other issues.

ary schemes are approximately actuarial. The public sector schemes are financed by all public employees at a rate of 2% of gross earnings, and the benefits offers a defined benefit level after consolidation with the National Insurance. This consolidation implies that people with low incomes, little public-sector experience and part-time workers subsidise the pensions of high-wage full-time employees who stay on in the public sector throughout their career. The benefit level is calculated from experience and the final salary. Going to one public sector employer to another does not interfere with the public sector insurance schemes, one is seamlessly transferred to the scheme of the new public employer as if all past public-sector contributions were to the last employer's fund. This provided the rationale for standardisation of public sector pension schemes (Norges offentlige utredninger 2000).

In Section 2 I present a model of sectoral choice and retirement in an environment where there is no aggregate uncertainty but where there are idiosyncratic shocks to earnings and human capital accumulation. The following sections present the data, explain how the model is estimated, the structural estimates and a policy experiments. An appendix provides additional details about the Norwegian retirement benefit system with some excessive detail, to make it possible to evaluate the relevance of the incentives embedded in the present model.

2 Modelling career choices

The full level of detail in the Norwegian system is impossible to embed in an empirical model of optimising individuals. I work with the simplifying assumption of no aggregate uncertainty, and abstract from the consolidation of married couples pensions. I do this by modelling the incentives for single men only, and apply these rules to men irrespective of their marital status. Since men tend to be older than their wives, have more work experience and higher wages, this is probably a reasonable first approximation to the marginal pension benefits. Modelling the interactions of spousal choices at the level of detail needed to model responses to policy changes in retirement benefits is currently not a tractable problem in life-cycle models, where households cannot be taken as stable (although important progress has been made by van der Klaauw and Wolpin (2005)). In the model, individuals have perfect foresight about changes to the benefit schemes.

I also attempt to produce the simplest possible model that can retain some of the incentives in human capital and pension benefits accumulation. This means that all variables will take on a discrete number of values, and many of them will be binary. Whereas it has been shown that complicated models can be estimated when appropriate approximations are made (French 2005; van der Klaauw and Wolpin 2005), this typically depends on the limiting restrictions of all unobservable variables having no history-dependence, and a strongly restricted structure on the unobserved permanent heterogeneity. I formulate a simple model, but one that allows for both unobserved differences in skill-technology and incomplete observation of the relevant state of individuals. The choices are, as in the baseline human capital papers, made to maximise the net present value of the future income stream. Hence there is no insurance motive to the retirement benefits in the model.

2.1 States and choices

Time is characterised by two indices, t index the age of the individual, $t \in \{0, \ldots, T\}$. To allow for more than one cohort when government policy and skill prices change over time, there is also an index $d \in \boldsymbol{D}$ that index different demographic groups, generations, that face different policy trajectories. Together, d and t determine the calendar year. The state of an individual, denoted by θ, is

$$\theta_t \equiv (\alpha_{t-1}, \mathrm{elig}_t, \boldsymbol{x}_t, \mathrm{pp}_t^f, \boldsymbol{s}_t, u_{4t}, \boldsymbol{\varepsilon}_{wt}, \mathrm{edu}_t). \tag{1}$$

I use α to denote the choice taken, hence α_{t-1} is the choice taken in the previous period. The "elig" variables index eligibility for National Service and early retirement. Experience, $\boldsymbol{x} \equiv (x_1, x, x^{NI})$, consists of three variables that represents experience in the public sector, x_1, total experience, x, and experience within the National Insurance system x^{NI} (needed to model those who started working before National Insurance was introduced). The capital within the National Insurance system is kept track of in pp^f, the final pension point score. Sector specific skills are held in $\boldsymbol{s} \equiv (s_1, s_2)$. The productivity in home production is u_4, and $\boldsymbol{\varepsilon}_w \equiv (\varepsilon_1, \varepsilon_2)$ is the realisation from the conditional wage distribution. While not relevant for individual choice, $\mathrm{edu}_t \in \{0, \ldots, 9\}$, counts past years of education over and above mandatory education (9 years).

Table 1: Possible and mutually exclusive actions

α	action
1	Work full time in the public sector
2	Work full time in the private sector
3	Go to school
4	household production
5	retirement

I will return to the exact specification of the the states, for now it is sufficient to note that an action α_t taken from the finite set $\boldsymbol{A}(\theta_t, t)$ of possible actions induce a distribution over next period states $\theta_{t+1} \in \boldsymbol{S}(\alpha_t, \theta_t)$, with probabilities

$$
\begin{aligned}
P(\theta_{t+1}|\alpha_t, \theta_t, t, d) = P_x(\boldsymbol{x}(\theta_{t+1})|\alpha_t, \boldsymbol{x}(\theta_t), t, d) \\
\cdot P_{\mathrm{pp}^f}(\mathrm{pp}^f(\theta_{t+1})|\alpha_t, w_{\alpha t}(\theta_t), \mathrm{pp}^f(\theta_t), t, d) \cdot P_s(\boldsymbol{s}(\theta_{t+1})|\alpha_t, \boldsymbol{s}(\theta_t)) \\
\cdot P_{u4}(u_{4,t+1}|u_{4,t}) \cdot P_\varepsilon(\boldsymbol{\varepsilon}_w) \cdot P(\theta_t). \quad (2)
\end{aligned}
$$

The experience sub-process depends only on choices and time, the retirement capital depends on choices, calendar time and the rest of the state only through the wage earned. The skill process depends only on choices, the productivity in home production depends only on the current level and the wage shock is independent of history and the other state variables.

The action set depends on the state because there are restrictions on when retirement is allowed, and some further restrictions will be imposed for computational reasons. There is also the possibility of the government restriction choices by forced and mandatory drafting into the National Service (I will return to this in section 3.3). In the first period, individuals are 16 years old and exiting compulsory education. In the last period T, only retirement is allowed and retirement brings the career to a full stop. The possible actions are enumerated in Table 1.

Experience is sector-specific and acquired by working. In order to limit the size of the state space, experience is acquired in step intervals of 10 years. In order to smooth the effect of these jumps, experience is acquired stochastically: By working one has probability 1/10 of a step increase in experience. Total experience is bounded at 40 years. The experience vector $\boldsymbol{x} = (x_1, x, x^{\mathrm{NI}}) \in \boldsymbol{X}$

keeps track of x_1, experience in the public sector and x, total experience, and

$$\boldsymbol{X} = \{(i,j,k) : i,j,k \in \{0,10,20,30,40\} \text{ and } i \leq j,\ k \leq j\}.$$

For everyone who entered the labour market after the introduction of the National Insurance, $x^{\mathrm{NI}} \equiv x$, and there are only 15 points on the grid in total. For the cohorts that made their early career choices without an existing National Insurance system in place, there are more possible points because total experience can be greater than the National Insurance experience. Limiting attention to those who entered ten years before 1967, I allow only the attainment of one unit (10 years) of experience earned before 1967, and this keeps the total number of experience points down to 29.

People leave compulsory schooling without any specific skills, but can acquire sector-specific skills by working or attending school. The specific skills take only two values (high and low), such that $\boldsymbol{s} \in \{0, \bar{s}_1\} \times \{0, \bar{s}_2\}$. Production of specific skills is not deterministic, but attending school there is a probability p_3^s that an unskilled individual acquires skills. Schooling produces general skills, by which I mean that public and private sectors skills jump simultaneously. There is a probability p_+^s that an unskilled individual working in sector will acquire a unit of specific skills in that sector and a probability p_-^s that a skill unit in a sector will be lost by an individual who does not use it.

The National Insurance administration calculates pension benefits conditional on a much larger state-space than can possibly be allowed for in a model to be estimated. The National Insurance keeps track of an "average pension points" which is the average over the 20 with highest acquired pension points. Pension points are calculated as a non-linear function of earnings, and the formula for calculating them has changed over time. Since 1992, for earnings below $1G_t$, one earns no pension points. For earnings in between $1G_t$ and $8G_t$, one earns $(y/G_t - 1)$ pension points, and above $8G_t$ there is no marginal contribution of earnings to pension points. Since 1992, the value of G_t is indexed to wages, and $G_{2000} = 48400$, approximately 7500 USD. See the appendix for more details on how this has been calculated historically.

In order to model the average-of-twenty-best years exactly, one would need to keep track of an pension point history as a point in $\mathbb{R}_+^{20}$. This is not feasible. The problem of keeping track of continuous variables can easily be solved by an appropriate discretisation. Keeping track of how the average over the twenty years of highest earnings evolve is a more difficult problem, and some approximation will have to be found. Researchers studying Social Security in the United States are faced with a similar problem.[3] French (2005) assumes that above average earnings always replace an average year after the history is full, while van der Klaauw and Wolpin (2005) assume that *any* current earnings always replace an average year. Since the current average will always be higher than the minimum of past earnings (the earnings point actually replaced under Social Security regulation), it would seem that these approaches systematically underestimate the effect earnings late in the career can have on retirement benefits. Rust and Phelan (1997) takes a different approach by estimating a reduced form regression where Social Security Averaged Indexed Monthly Earnings (AIME) depends on current earnings, AIME lagged by one period and age.

Since the previous public sector wage can be predicted by the state-space, I do not track it separately. I follow the method of Rust and Phelan (1997), but without treating the National Insurance capital as a continuous variable. I discretise the accumulated pension average into four levels, and estimate reduced form logit-equations for whether a given earnings level give rise to a jump on the grid, such that the distribution of next period accumulated pension average is a continuous function of current earnings. I will return to the details of this procedure in Section 3.5 after having described the available data.

2.2 Utility of choices

Based on their outcome (θ_t, α_t) individuals receive a per period utility of $U(\theta_t, \alpha_t)$:

$$U(\theta_t, 1) = (1 - \tau_{dt}(w_{1dt}(\theta_t)) - 0.02) w_{1dt}(\theta_t) + u_1, \tag{3a}$$

$$U(\theta_t, 2) = (1 - \tau_{dt}(w_{1dt}(\theta_t))) w_{2dt}(\theta_t), \tag{3b}$$

$$U(\theta_t, 3) = u_3 - \mathcal{B}(\alpha_{t-1} \neq 3) u_{3r}, \tag{3c}$$

$$U(\theta_t, 4) = (1 + \xi)^{t-d} u_{4t}, \tag{3d}$$

$$U(\theta_t, 5) = B(\theta, t, d). \tag{3e}$$

[3]The US Social Security Administration keeps track of 35 years of highest past earnings in the "Average Indexed Monthly Earnings".

Wages are functions of the state, and τ_{dt} is the average tax function. In the public sector, workers also contribute 2% towards future retirement benefits. All agents are risk-neutral. If they work they receive a wage payment and potentially a non-pecuniary value u_1 of working in the public sector. The net present value of the pension benefits are $B(\theta, t, d)$, and this function allows for taxation, mortality and the details of the benefit systems (outlined in Appendix A). If they go to school there is a cost (in money-terms) of $-u_3$ and possibly an extra cost u_{3r} of returning to school from a working or a household production state.[4] Productivity in home production is u_{4t}, this also increases at the rate of technological progress.

Wages are functions of experience, current skills and a conditional wage distribution with discrete support,

$$w_{tj}(\theta) = (1+\xi)^{t-d} r_j \exp\left(s_j + \beta_{j1} x_j + \beta_{j2} x_j^2 + \beta_{j3} x_{i(\neq j)} + \varepsilon_j\right).$$

The skill-price is determined outside of the model, but there is (potentially) secular growth in labour productivity. The stochastic element ε is distributed on a discrete support with a known covariance-matrix. Note that whereas education produces skills with some probability, education is not in itself a determinant of wages. There is also a technologically neutral growth ξ in labour productivity.

Any attempt at modelling the value of retirement will be flawed by an inaccurate modelling of the intricate nature of both the legal code and of the description of a state necessary. In this paper I model the total gross flow of benefits as the maximum of the potential National Insurance, b_{dt}^{NI} and the public sector benefits, b_{dt}^{gov}, and define

$$b_{dt}^{\text{gross}} = \max\{b_{dt}^{NI}(\theta_t), b_{dt}^{\text{gov}}(\theta_t)\}. \tag{4}$$

The *net* benefits are

$$b_{dt}(\theta_t) = \mathcal{B}(t<T)u_4(\theta_t) + \mathcal{B}(\text{elig}(\theta_t)=5)\left[b_{dt}^{\text{tax}} + (1-\tau_{dt}(b_{dt}^{\text{gross}}))b_{dt}^{\text{gross}}\right]. \tag{5}$$

Home production and retirement can be combined if $t < T$, and the tax-system favours old-age pensioners in several ways that are partially off-setting. The net effect of this does not vary much with income (Rønningen and Fredriksen 2002) and can hence be modelled by a constant b_{dt}^{tax}. If the individual is eligible for

[4]The function $\mathcal{B}(\psi)$ is an indicator function that takes the value 1 if ψ is true and 0 otherwise.

retirement, he gets the tax-reduction and benefits net of taxes.

The yearly National Insurance component of retirement benefits for a single male is approximated by

$$b_{dt}^{NI} = G_t \max\{1.79, 0.42\mathrm{pp}_t^f \cdot x_t^{NI}/40\}. \tag{6}$$

The public sector component is approximated by

$$b_{dt}^{\mathrm{gov}} = \begin{cases} 0.67 \cdot \min\{x_{1t}/30, 1\} \cdot \min\{w_{rt}, 12G_t\} & \text{if retirement from public sector} \\ 0.67 \cdot \min\{x_{1t}/40, 1\} \cdot \min\{\tilde{w}_{1t}, 12G_t\}, & \text{else,} \end{cases} \tag{7}$$

where w_1 is the wage at the time of retirement, x_1 is the experience in the public sector and $\tilde{w}_1$ is the last recorded wage. Since $\tilde{w}_1$ is not kept track of separately in the state-space, I use the current state to predict what the last public sector wage would have been. I will make the restriction on choices that no inter-sectoral career changes can be made after the age of 60 I therefore predict public sector wages of those not retiring from the public sector back to the this age netting out the effect of technological progress.

Everyone is eligible for retirement when they reach 67. Everyone working in the public sector are eligible for early retirement, "AFP" when they reach 62. In the private sector, only about 50% of the firms are covered by the programme.[5] Without modelling firm-to-firm transitions in the private sector, only a crude approximation to private sector eligibility can be made. At the age of 60, I randomly let 50% of the private sector employment be eligible for the AFP. The AFP programme itself I approximate by an earlier possible retirement age, but keep the rest of the regulations from the ordinary retirement. See Hernæs et al. (2000) for a more focused analysis of the AFP programme.

Having defined the flow from retirement, the net present value at a given moment in time t, discounted by a discount factor δ and the net survival probabilities, with m_j representing the mortality rate, can then be written

$$B(\theta_t, t, d) = \sum_{i=t}^{95} \delta^{(i-t)} b_{di}(\theta_t) \prod_{j=t}^{i} (1 - m_j). \tag{8}$$

[5]The exact number is disputed (Midtsundstad 2004).

The choice of terminal age at 95 is not important.

2.3 value of choices

The total value of an outcome is affected by the state one is at and the choice one makes, but for expositional clarity it is convenient to introduce $p^m(\theta_{t-1}, \alpha_{t-1}, t, d)$ which is the probability with which the government will enrol an individual in the national service in period t. This could potentially depend, on lagged outcomes, such that individuals can control the draft to some extent.[6] If the National Service is forced on individuals, their possible next-period states are $\theta' \in \boldsymbol{M}(\theta)$. The transition to a new state is governed by the Markov transition probabilities $P(\theta'|\alpha, \theta, t, d)$. This depends on age, which determines eligibility for national service, and calendar time which reflects changes in the retirement benefit system – but also the effects of increasing wages on the probability of a transition on the final pension points score. In my application, and for computational reasons, I will only update the transition matrix every fifth year. I restrict the National Service to be equivalent to the possible home-production states, $\boldsymbol{M}(\theta) = \boldsymbol{S}(4, \theta)$. There is then an expected value $v(\alpha, \theta)$ associated with every choice,

$$v_t(\alpha, \theta) = U(\alpha, \theta) + \delta(1 - m_t)\Big[p^m(\alpha, \theta, t+1, d) \sum_{\theta' \in \boldsymbol{M}(\theta)} \frac{P(\theta'|4, \theta, t, d)}{P(\boldsymbol{M}(\theta))} V_{t+1}(\theta') + (1 - p^m(\alpha, \theta, t+1, d)) \sum_{\theta' \in \boldsymbol{S}(\alpha, \theta)} P(\theta'|\alpha, \theta, t, d) V_{t+1}(\theta')\Big]. \quad (9)$$

The state is fully observed before the choice is made, so the value function V_t is the maximum of the values of the possible actions,

$$V_t(\theta) = \max_{\alpha \in \boldsymbol{A}(\theta, t)} v_t(\alpha, \theta). \quad (10)$$

The current action determines the set $\boldsymbol{S}(\alpha, \theta)$ of reachable next-period states. If the current action was to retire or the individual reaches mandatory retirement age, $t = T$, this set is empty and the career ends.

[6]This is similar to how Lee (2005) models the accumulation of children. I am, however, not letting choices affect the probability of being drafted since we will find in the next section that the average transition probabilities does not reflect the selection mechanism most people would expect.

We must allow for some smoothing of the choice probabilities, or there would be a degenerate choice distribution for every θ. This would a lead to a non-smooth objective function in the estimation procedure. Following Ferrall (2002), I therefore impose a ρ smoothing of choices, with $\rho \in [0, 1)$,

$$\tilde{v}_t(\alpha, \theta) = \mathcal{B}(\alpha \in \boldsymbol{A}(\theta)) \exp\left(\frac{\rho}{1-\rho}(v_t(\alpha, \theta) - V_t(\theta))\right), \tag{11}$$

and let individuals have exponentially smoothed choice probabilities

$$P_\alpha(\theta, t, d) = \frac{\tilde{v}_t(\alpha, \theta)}{\sum_{\alpha' \in \boldsymbol{A}(\theta, t)} \tilde{v}_t(\alpha', \theta)}. \tag{12}$$

The extreme case of $\rho = 0$ is the "perfect smoothing" case with all choices being given equal probability. As ρ approaches 1 the choices approach the degenerate case with $\arg\max_\alpha v(\alpha, \theta)$ being chosen with probability one. The value of ρ is estimated. In practice, and with regard to predicted choice probabilities, this method is not much different from following the tradition of adding extreme-value added terms to each choice.[7]

Having solved the backward induction for $P_\alpha(\theta, t, d)$ given $P(\theta'|\theta, \alpha, t, d)$, one can solve forward for the full distribution of states at any time, let $\mu(\theta, t, d)$ be the probability mass of generation d in state θ at age t,

$$\begin{aligned}\mu(\theta, t, g) = \sum_{\theta'} \sum_{\alpha} \Bigg\{ \Bigg[& p^m(\alpha, \theta', t, d) \frac{\mathcal{B}(\theta \in \boldsymbol{M}(\theta'))}{P(\boldsymbol{M}(\theta'))} P(\theta | 4, \theta', t-1, d) \\ & + (1 - p^m(\alpha, \theta', t, d)) P(\theta | \alpha, \theta', t, d) \Bigg] P_\alpha(\theta', t-1, d) \mu(\theta', t-1, d) \Bigg\}. \end{aligned} \tag{13}$$

With the assumption that individuals start after leaving middle school without skills and experience, this defines a full distribution of states that can be compared with data. I will return to the empirical application in Section 4.

[7]Although it can be argued that the method I employ have advantages in situations where the size of the choice set changes over the life-cycle.

2.4 Computational restrictions on choices

For computational reasons, I restrict the possible choice of attending school to the first 25 years after leaving middle-school. In order to fit the pension system, I will not let anyone retire (in the sense of claiming the National Insurance and the supplementary benefits) before they are 62 years old. There are few people changing their career from the private sector back to the public sector in the last years of their career, although it might seem as if this would be favourable for those with previous experience in the public sector. Undoubtedly there are problems with switching career the year before retirement that will be difficult to capture in this model, so I will not allow anyone to switch sectors after the age of 60.

3 Data and ancillary models

Statistics Norway has collected data from various government administrative records, from which the labour economics group at Norwegian School of Economics and Business Administration has constructed an integrated database (Møen, Salvanes and Sørensen 2003). The entire population of residents in Norway aged 16–74 can be followed with yearly observations in the years 1986–2000. In addition, a full panel of yearly labour earnings since the introduction of the National Insurance scheme in 1967. Labour earnings is measured using the National Insurance definition, which includes benefits while sick but excludes unemployment benefits.

For the purpose of this paper, since there is no immigration or emigration in the model, I restrict attention to those born in Norway (or resident at the age of 16 when the choice-model begins). I use three cohorts for estimation, the 1941, 1955 and the 1970 cohort. The 1970 cohort is a natural choice since these people leave compulsory schooling in the first year from which we have data. The 1955 cohort is chosen to have a middle-aged group, and the 1941 cohort since this generation can accumulate exactly 10 years of experience before the introduction of the National Insurance scheme. The model assumes an educational environment that is constant. This is an abstraction, an educational reform in the 1960's meant that compulsory education for the 1941 cohort in fact ended earlier, and that not everyone stayed on in school until they were 16 (Raaum, Salvanes and

Sørensen 2003). Considering the time in between schooling and observation, I assume this to have only minor impact on the estimates. The population in the oldest cohort is, at the time of our observations, 20498 persons, while the 1955 and 1970 cohorts consists of 30670 and 32422 respectively.

The attachment of workers to firms is observed at a point in time, from 1986 to 1994 this is May 31st, from 1995 and onward this observation is made in November. For each of these attachments we get to observe the start-date of employment and, if applicable, the termination date. So for choices, the calendar year is not the only possible choice of yearly time-period. For any study that includes education it is more convenient to measure the year according to the academic year. I have defined the choice period of a year as beginning July 1.

Any model of career choices will come up against the problem that in the data we see people combining several different choices within the defined time period. People may take on a part-time job while still in school, some people may quit school in the middle of the year to start work, some drop out without any evident cause, and others take on a succession of high-frequency jobs. Any rule that attempts to force data into year-long decision periods and exclusionary choices will be arbitrary to some extent. I follow Keane and Wolpin (1997) and Lee (2005) in hierarchically classifying the individual into categories, starting with education, following with the two working choices and at last let the home-production category be defined residually.

3.1 Defining choices from data

Education For education, we get to observe the highest education level achieved by individuals, and the graduation date of this level. From 1986 and onward, we also get to observe the flow of education from administrative records; whether an individual was enrolled in part-time or full-time education as of November 1st (Vassenden 1995; Statistisk sentralbyrå 2001). I use the "enrolled in full-time education" to classify the individual as enrolled in school. Possibly this could be refined by checking to see that the individual actually graduates, but this is complicated by the fact that some educational categories span several years, and a not insignificant fraction of the students seem to move laterally in the grade system in a way that is not captured well by the stock variable of "highest level achieved". Education is measured as the number of years above mandatory

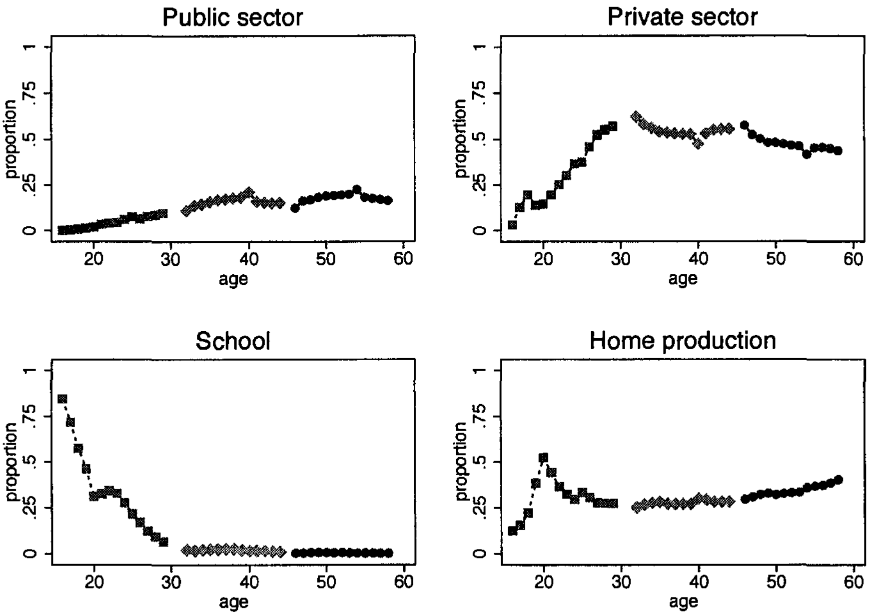


Figure 1: Choice distributions. Classifications are defined in the text. Three cohorts, the 1941, then 1955 and the 1970 cohorts are shown, with choices recorded in the window 1986–2000.

education, and is capped at 9 years.

Work As mentioned, we observe if the individual is attached to an establishment once a year. We also get to observe ownership-structure, industry, a categorical measure of weekly hours, and start and stop dates.[8] Using this history of work-spells, I calculate the hours worked in the year as defined from July 1 to June 30 the following year using the Labour Force Survey as guide to how many hours per week to impute within the categorical groups, and I use only the job spells of 20 or more hours per week.[9] Since we only have yearly snapshots of workplace attachment, this method may not work well when there are a lot of high-frequency job changes. I have therefore chosen a cut-off for whether an individual is classified as working as low as 1000 hours per year. In order to qualify as "working" the individual also needs a minimum level of earnings. Since the working / academic year spans two calendar year, I demand that in at least one of those two calendar years the individual earned $1G$ (with G defined as by the National Insurance Administration). If they in fact worked that year, this should easily be fulfilled, but some workplaces may wrongfully keep people in their records after short summer jobs, and if individual do not attach themselves to other workplaces or go to school, this could cause errors in the classification.

Sectoral categorisation For those categorised as working following the above definition, I make the categorisation into public and private sector based on the job-spell that contributed most hours. We get to observe a unique identifier of the establishment, and this identifier can be used to link to the central register of establishments and corporations (Olsen 1993), in which there is a variable indicating whether the establishment is owned by central or local government. When this link works, this is the sectoral categorisation. I have only had access to the data from the central register of establishments and corporations for the period 1986–1995. For the remaining years 1996–2000, I have used the public/private categorisation as available in 1995 for all those establishments that existed then. For new establishments, I have used the 1995 data to create a mapping from industry code (ISIC Rev2 at the 4-digit level) to sector (using the within industry mode of the "public" indicator, weighting with the number of employees) and

[8]Weekly hours is categorised in 3 levels, (1) less than 20 hours per week, (2) between 20 and 30 hours and (3) more than 30 hours per week.

[9]I have calculated averages of total hours for men over the period 88-95, these averages are 12.59, 25.65 and 36.84 hours per week.

used this industry-based categorisation on the new establishments 1996 to 2000. For the years 1999 and 2000 a new mapping from the ISIC to the NACE industry code is needed, I use the within NACE mode of the ISIC codes, again weighted by the number of employees and at the 4-digit level.

Residual home production and National Service Those not categorised under the above rules are classified as at home, an observational group that also includes the (observationally equivalent) individuals in the National Service.

The final distributions of choices as people get older for the 3 cohorts used in estimation is shown in Figure 1. As we can see, there is slow but steady growth in public sector employment as people get older, whereas private sector employment peaks around the age of 30, and then slowly tapers off as people get older. There are few people in education after the age of 30, and we see that the slow decline in private sector employment is (nearly) matched by a slow growth of people in home production. The early spike in home production reflects that this category also contain those in National Service.

In Figure 2 we see the average conditional choice probabilities by age. There is, of course, a lot of noise in the transition probalities from the choices few make. We can, however, see that there is a quite a lot of persistence in choices and that there is more mobility among the young than among the old.

3.2 Wages and earnings

Because I define the decision year as starting at the beginning of the third quarter, calculating wages is difficult, since the data I have to work with is yearly earnings, reported for the calendar year. However, since I do not attempt to match individual level data, I will not have to match individual decisions to individual choices. It is therefore possible to take some liberties when calculating wages. I have in fact chosen to offset the measurement of wages by half a year, such that the calendar years wages corresponds to the decision period that starts in the third quarter of that year. Since the square of wages is used when matching the model to the data, measurement error in wages/earnings will have consequences for identification. I have therefore chosen to only use the earnings of full-time workers who can be matched to a workplace for at least 320 days of the relevant calendar year. I have then taken the means of the labour earnings at 1998 consumer prices as the measure of that period's full-time return to labour. All

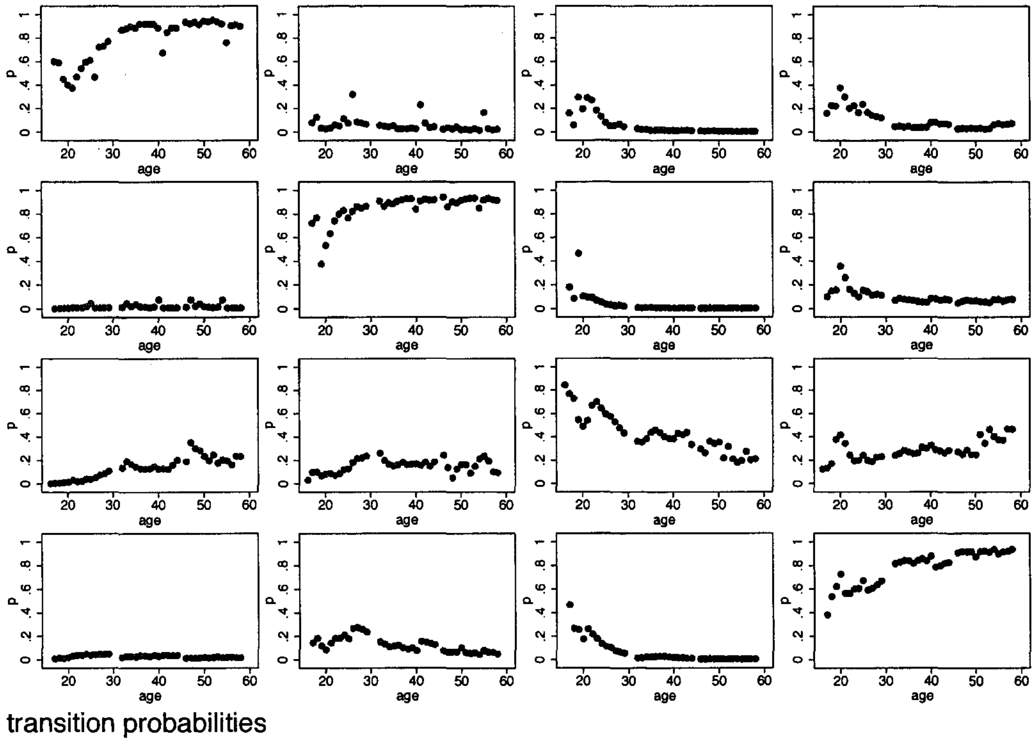


Figure 2: Choices conditional on previous choices.

Note: A $p_{ij}(t)$ entry is the fraction of people who choose i in the previous period who choose j at age t. The i and j's range from (1) work in the public sector, (2) work in the private sector, (3) attend school, to (4) stay home (observationally equivalent to National Service). Calculated on the 1941, 1955 and 1970 cohort in the observational window 1986-2000.

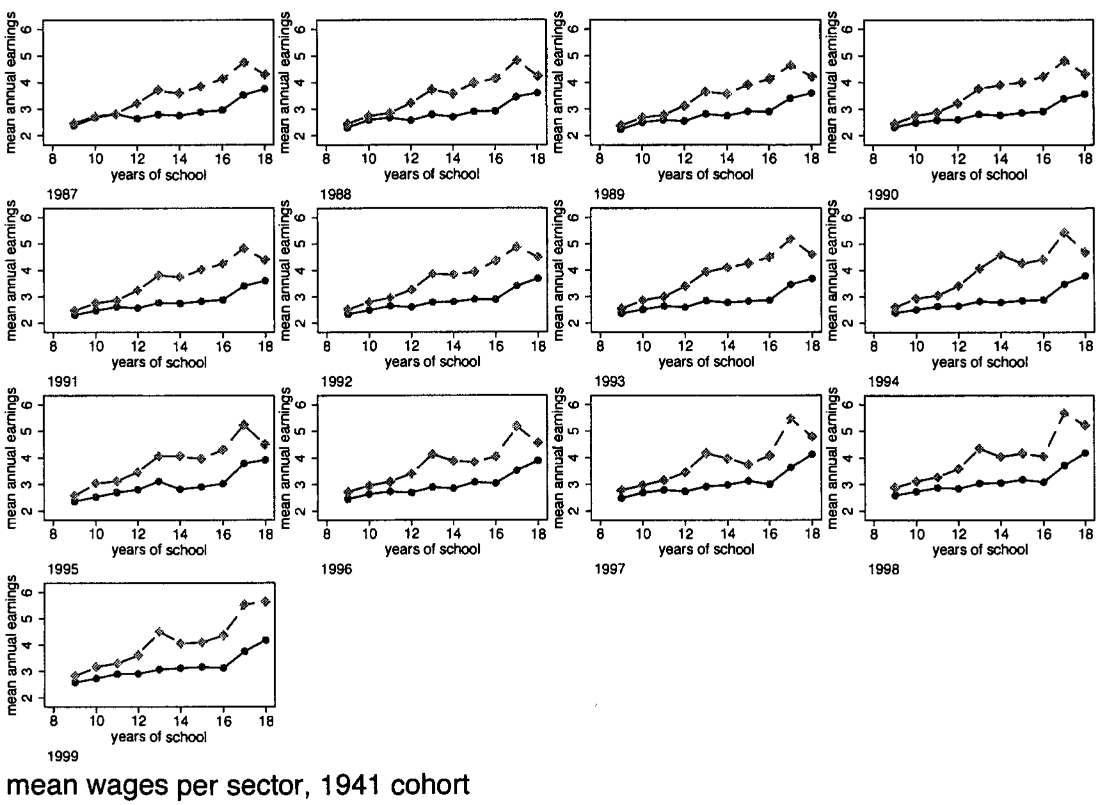


Figure 3: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

money measures throughout the paper are calculated in terms of 1998 NOK.

The figures 3–5 shows the earnings/education profiles in both the public and the private sector for the years the model is estimated on. As we can see, among the older cohorts, there is a substantial gap between private and public sector wages, and a clear positive earnings profile with respect to education. In the younger 1970 cohort, we see neither as clearly. It seems as if experience can partially offset education in the first years of the career, and that the public sector has a flatter experience profile than the private sector.

Since the private sector supplementary pension benefits are transferrable, and the variation in types of private supplementary benefits is too large to take into account, I impute pension premiums to current wages in the private sector. I use

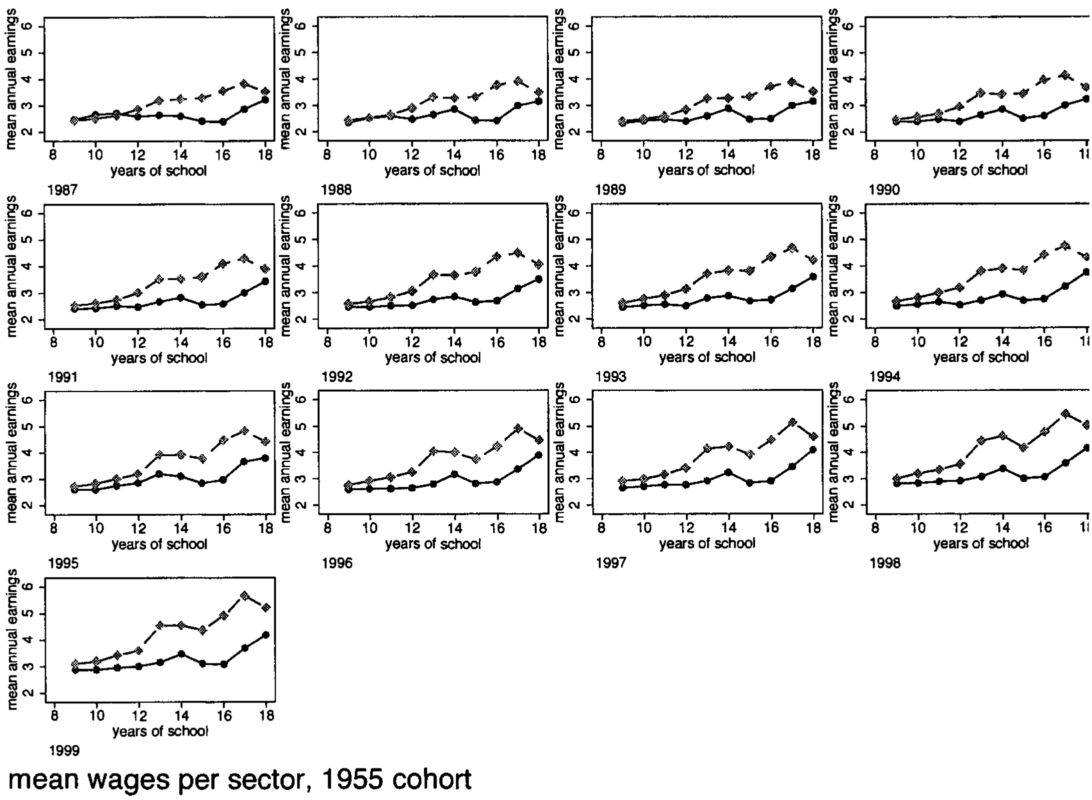


Figure 4: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

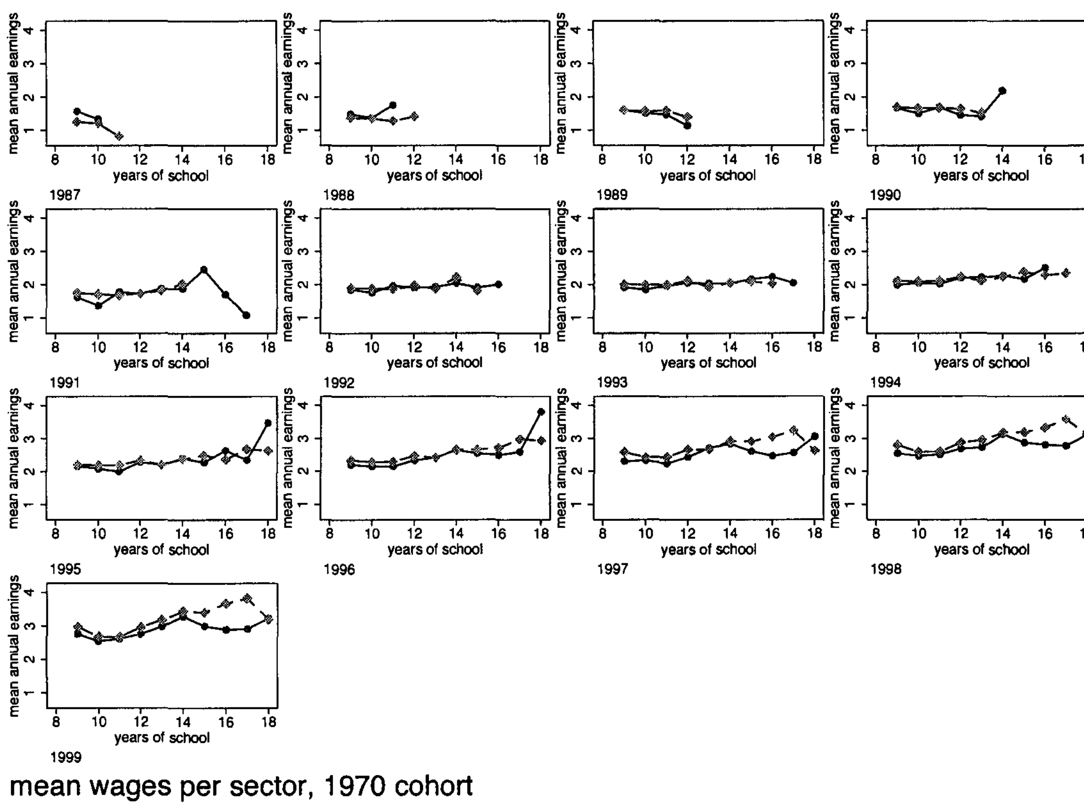


Figure 5: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

the coverage rates from Pedersen (2000) and the rates I have shown in Figure 10.

3.3 National Service

In Norway most young males serve a year of military services. This is different from the military career-choice modelled by Keane and Wolpin (1997) in that it results from a mandatory military service for all young males. In principle, everyone is called to serve for a year at age 19, but in fact there are a number of exceptions. First, not all are in fact called to serve, some are turned away because of weak health or because the full cohort is not needed. Second, it is possible to postpone military service if it would interfere with studies, and third; one can refuse military service as a conscientious objector, in which case there is mandatory civilian service.

The military service is also only partly documented in the data available to us (being only observable from 1993 and onward). This makes difficult to model the early career choices of young men, since military service will potentially corrupt the first 5 or 6 years of young men's career histories.

These fact that military service is only partially controlled by individuals also make it difficult to model the interruption to young men's careers as either completely exogenous or freely chosen by individuals. An intermediate solution, in which eligible young men who find themselves called to serve with probability $p_{tg}^{m}(\alpha_{t-1})$ is possible, by letting the National Service arrive at a rate that depends on the choice in the previous period it would be possible to avoid military service by staying in school hoping that one will not be called up. From Table 2 it is evident that it attending school is not a safeguard against being drafted. In fact, draft-probabilities are higher among those who attended school the previous period than among those who worked or stayed home, a somewhat surprising result. A full understanding of this will probably require a more detailed modelling of the selection into National Service, a task the model in this paper is not fit for. As a first approximation, I let the probability of being drafted depend only on age and cohort.

Since individual level information on National Service is only available from 1993 and onward, most of those serving will wrongly be classified as engaged in home production. In order to extend the information available, I have collected aggregate historical information from the National Draft Board about how many

Table 2: Probability of being drafted for National Service, $p^{m}_{1975,t}$

year	Choice in previous period			
	work	school	home	total
1993	0.0520	0.0204	0.0564	0.0239
1994	0.1751	0.2476	0.1637	0.2351
1995	0.2770	0.2992	0.2899	0.2978
1996	0.1633	0.2347	0.1602	0.1993
1997	0.0609	0.1520	0.0628	0.0960
1998	0.0440	0.1234	0.0468	0.0560
1999	0.0104	0.0356	0.0154	0.0016

Note: Calculated on the 1975 cohort. Draft is defined as the first entry into the military or the social service that lasts more than 100 days.

of a given cohort they have in fact trained. I then propose a multiplicative model of draft intensity, where

$$p^{m}_{dt} = \zeta_d \cdot p^{m}_{1975,t}. \tag{14}$$

The empirical hazard rate for the 1975 cohort is available in Table 2, and I then scale these hazard rates by a cohort-specific intensity ζ_d that is calibrated so that the process fits the aggregate proportion of a cohort that has been trained.[10]

The draft into the National Service is mostly in the beginning of every quarter. The third quarter has the largest numbers of draftees, so that for most, conscription is more or less synchronised with the academic year. Since most draftees spend about 9 months in the service, I want to classify them as drafted the decision year they spend to the largest extent in the National Service, so I categorise those who enter between April 1 this year and March 31 the following year to be in the National Service in the decision year that starts July 1 that year. In order to define the entry date, I look at the first spell that lasts more than 100 days.[11]

[10]Numbers on cohort specific training data have been provided by the National Draft Board ("Vernepliktsverket").

[11]After first joining smaller sub-spells with less than a month in between them, as these are mostly transfers from one army base to another.

3.4 Taxes

Since individuals aim to maximise the present value of earnings and utility flows net of taxes, I need the tax-functions τ_{dt} for the years 1956–2040. I have constructed the income tax-functions for the case of an individually assessed individual with no wealth or capital income for the years 1956–2000. My income-tax measure includes local and central government taxation, temporary income taxes and reliefs and National Insurance premiums.[12] These functions turn out to be so complex and non-smooth that they cannot be used directly in estimation. Instead I have fitted fourth-degree polynomials to the average tax-rate, using the distribution of yearly earnings of those 40 years old as weights. Since I do not have annual earnings for the 1956–1966 years, I have scaled the 1967 earnings back using the index of manufacturing wages (table 10.3, Statistisk sentralbyrå 1995). This sequence of polynomials fit within sample fairly well. Out of sample, particularly for high earnings, these polynomials do not fit well. I therefore use the distribution of earnings to calculate the 90th percentile, above which I restrict the average tax function to be flat.

Since the choices of young individuals in the 1986–2000 period depend on their projections about future taxation, any choice about how to calculate future taxes is open to debate. I have chosen to use the polynomial that approximates the 2000 tax function on the 2000 distribution of earnings. In an attempt to construct the future taxes that keeps the current rate of redistribution more or less constant, I scale the parameters from this polynomial while estimating, using the estimated rate of labour-augmenting technical progress to stretch the tax function.

Pension benefits are treated favourably by the tax system: a lower rate is used to calculate National Insurance premiums, there is a special allowance for old age and there are rules limiting the total taxes of pensioners. These rules partially offset each other, and Rønningen and Fredriksen (2002) show that the total effect lies in a fairly narrow band, it varies from 9900 to 15700 (both in NOK98) in 1995 (Table 4 Rønningen and Fredriksen 2002). I use the midpoint of this, NOK 12800, as an estimate of the net tax benefit of retirement in 1995, and adjust this value at the same rate as the basic pension unit – at the rate of technological progress estimated.

[12] I have drawn on Statistisk sentralbyrå (1958, 1968, 1975, 1988, 1994) and current numbers from `http://www.odin.dep.no/`.

3.5 Accumulated average of past pension points

I use the actual distribution of final pension points of those aged 60-70 years old in 2000 to calculate a 4-point grid on the final pension points, at the 12.5%, 37.5%, 62.5% and 88.5% percentiles, (3.03, 4.44, 5.50, 7.06). See Figure 6 for the full distribution of the final pension point score. Note that the 37.5th percentile is just enough to reach above the minumum pension for a single man (for which 4.26G in accumulated average is needed). It is then straightforward to classify individuals as being in the corresponding intervals defined by the 0-25%, 25%-50%, 50%-75% or 75%-100% percentiles (with cut-offs (3.90, 4.96, 6.18). For this truncation of the average accumulated pension points, I estimate reduced form logit-equations of the probability of jumping up or down on this grid, with the current pension points (a function of earnings, see discussion in Section A.1) affecting the probability of switching up or down. This way, the effect of current earnings on accumulated average pension points is smooth and continuous.

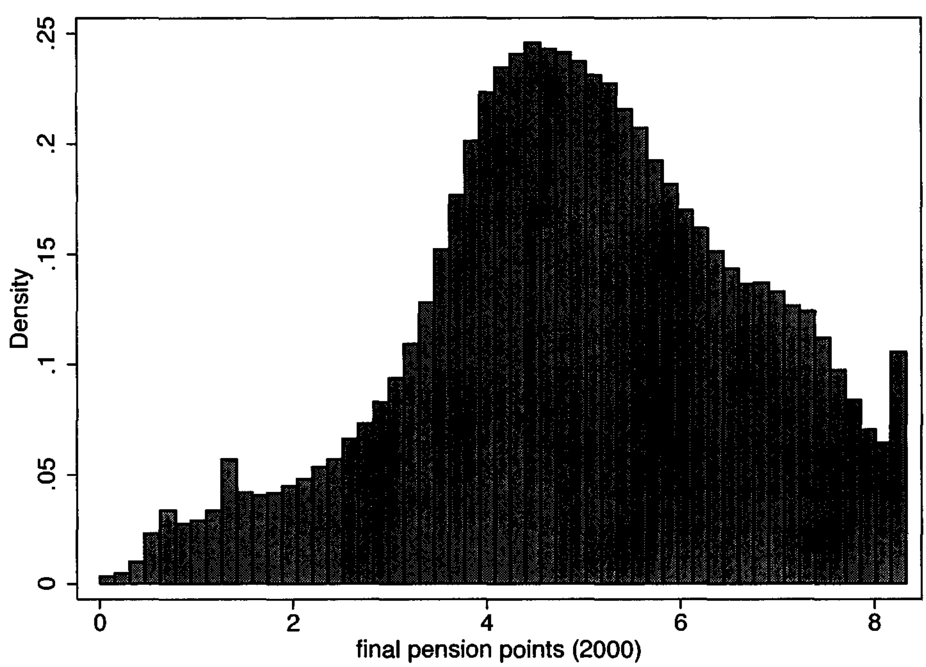


Figure 6: Final pension point score of potential retirees (60–70 years old in 2000)

Table 3: Average empirical transition rates on the pension capital grid grouped by experience. Calculated on the 1940, 1955 and 1970 cohorts.

	pp^f_{t+1}								
	0-9 yrs of experience					10-19 yrs of experience			
pp^f_t	1	2	3	4		1	2	3	4
1	0.972	0.027	0.000	0.000	1	0.951	0.049	0.000	0.000
2	0.061	0.840	0.092	0.000	2	0.013	0.930	0.056	0.000
3	0.005	0.070	0.851	0.075	3	0.000	0.015	0.952	0.034
4	0.002	0.006	0.072	0.920	4	0.000	0.000	0.017	0.984
	20-29 yrs of experience					30+ yrs of experience			
	1	2	3	4		1	2	3	4
1	0.912	0.084	0.000	0.000	1	0.968	0.032	0.000	0.000
2	0.000	0.916	0.084	0.000	2	0.000	0.979	0.021	0.000
3	0.000	0.000	0.945	0.055	3	0.000	0.000	0.992	0.008
4	0.000	0.000	0.000	1.000	4	0.000	0.000	0.000	1.000

Whereas one would not need more than a two-point discretisation of the accumulated pension points to capture the main incentives, a finer grid is needed in order to robustly estimate the reduced form process on the grid. Separate regressions are made for each experience category, with experience calculated as number of years with earnings of at least one G (this corresponds to National Insurance practise). The average transition rates by experience group is shown in Table 3. We see that from year to year, there is not excessive volatility on this grid. In Table 4 are the estimated logit-equations where the probability of a jump up (or down) on the grid is estimated as a function of current earnings (in pension points), with different parameters by experience group. These estimates are then applied as parameters in the structural model.

4 Estimation

In order to facilitate estimation, I introduce some further notation. Remembering that $\boldsymbol{D}$ is an index set of observable (demographical) groups, one such group could be all males born in cohort 1970. Within each of these observable groups, let $k \in \{1, \ldots, K\}$ index a set of unobservable types in the sense of Heckman

Table 4: Reduced form (Logit) estimates of final pension points process on the 4-point grid. Standard errors in parentheses. Estimated on data from the 1940, 1955 and 1975 cohorts.

	0-9 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up	down	up	down	up	down
pension points	1.62	−1.31	1.97	−1.28	1.94	−1.21
	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)
constant	−11.98	3.70	−15.03	4.93	−16.58	5.65
	(0.06)	(0.07)	(0.15)	(0.11)	(0.30)	(0.19)

	10-19 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up	down	up	down	up	down
pension points	1.31	−0.05	1.54	−0.75	2.00	−0.63
	(0.01)	(0.01)	(0.02)	(0.01)	(0.04)	(0.02)
constant	−10.51	−0.11	−13.26	0.19	−18.29	0.06
	(0.05)	(0.04)	(0.12)	(0.06)	(0.31)	(0.09)

	20-29 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up		up		up	
pension points	1.21		1.23		1.37	
	(0.01)		(0.01)		(0.02)	
constant	−8.25		−9.43		−11.72	
	(0.07)		(0.08)		(0.13)	

	30+ years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up		up		up	
pension points	2.45		2.76		3.02	
	(0.13)		(0.12)		(0.14)	
constant	−14.50		−18.03		−22.43	
	(0.68)		(0.63)		(0.90)	

and Singer (1984). Each of these K are associated with a weight $\lambda_{dk} \in [0, 1]$ (such that $\sum_k \lambda_{dk} = 1$) and a type-specific parameter-vector γ_k, the structural parameters. Now the data will be composed of DK types each solving their own problem as defined in the model described in Section 2, and from now on I will on occasion subscript the terms defined earlier by their dk group.

4.1 Measurements and moments

The distribution μ in equation (13) is defined as the distribution of individuals of individuals over the state space $\boldsymbol{\Theta}$. In itself, this is not useful for estimation since there are several unobservable components of this space. Let us therefore determine a measurement function $y : \boldsymbol{\Theta} \times \boldsymbol{A} \to \boldsymbol{Y}$, where $\boldsymbol{Y}$ is a space of measurements: yearly earnings, schooling, pension capital and choices in the previous period. Note that there are also endogenous state-variable that are not measured, such as sector-specific experience (for the older generations) and sector-specific skills, and choices can also be obscured by forced National Service. In the terminology of Ferrall (2002), the social environment is not transparent.

The distribution over states induced by the optimal choices defines a distribution function $F(\mu) : \boldsymbol{Y} \to [0, 1]$ on the measurements. With this distribution F it is easy to define a series of observable moments predicted by the model, this distribution $F_d(\boldsymbol{\gamma}, \boldsymbol{\lambda}_g)$ will depend on all the type-specific parameters ($\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$) and the population shares of the types in the d group ($\boldsymbol{\lambda}_d = (\lambda_{d1}, \ldots, \lambda_{dK})$). It is now straight-forward to define a moment-generating function $\Delta : \boldsymbol{\Gamma} \to \mathbb{R}^Q$, with Q being the number of moments.

The empirical moments I match are outlined in Table 5. This list is a variation of those used by Lee (2005), but with more conditional choice probabilities than Lee was able to calculate from his data. The reason for choosing to fit the mean and the mean squared wages instead of the mean and the variance such as Lee does, is that only one pass through the states is necessary. This makes an difference for computational efficiency since the problem for each of the K types can be calculated separately by parallel processors, and fewer numbers need to be passed along. In order to calculate the variance of wages one would need to pass the full distribution across states for all types. This reduction in the information that has to be passed between processes reduces the time needed for communication on inexpensive clusters with distributed and limited memory, and makes it possible to apply the efficient methods for finite mixture models developed by Ferrall (2005).

The model is estimated with GMM, with the objective function

$$\sum_{d\in \boldsymbol{D}} \left[\left((\widehat{\boldsymbol{\Delta}}_d - \Psi_{k=1}^{K} \left(\boldsymbol{\Delta}_d(\gamma_k), \lambda_k \right)) \right)' \boldsymbol{\Omega}_d \left((\widehat{\boldsymbol{\Delta}}_d - \Psi_{k=1}^{K} \left(\boldsymbol{\Delta}_d(\gamma_k), \lambda_k \right)) \right) \right]. \tag{15}$$

The Ψ operator aggregates the moments from the K unobserved types using the λ_k weights and ancillary information about conditional masses. The standard formula for the variance of the estimates, is now

$$\mathrm{var}(\widehat{\gamma} - \gamma_0) = (\boldsymbol{J}'\boldsymbol{\Omega}\boldsymbol{J})^{-1}\boldsymbol{J}'\boldsymbol{\Omega}\left(\mathrm{var}(\widehat{\boldsymbol{\Delta}})\right)\boldsymbol{\Omega}\boldsymbol{J}(\boldsymbol{J}'\boldsymbol{\Omega}\boldsymbol{J})^{-1}, \tag{16}$$

where $\boldsymbol{J}$ is the Jacobian of the moment vector with respect to the parameter vector.

Since the number of moments used is large I do not attempt an optimal weighting matrix. For computational tractability I follow use a diagonal weighting matrix outlined in Table 5. The estimation, like that of Lee (2005) and Lee and Wolpin (2004), is not fully standard in that conditional moments are matched (such as sector and education-level specific wages).[13] This means that the full covariance matrix of $\boldsymbol{\Delta}$ is not defined. As a practical approximation I use only a diagonal estimate of $\mathrm{var}(\widehat{\boldsymbol{\Delta}})$, calculated with the size of the relevant sub-populations.

4.2 Estimated parameters

Estimation was performed on the 5-node cluster "paulus" at Norwegian School of Economics and Business Administration using the FmOpt library (Ferrall 2005).[14] Tables 6, 7 and 8 contain the estimates of the structural parameters where only base levels and the sector specific skills are allowed to differ by type. The discount factor is fixed at 0.98, the same level as in Rust and Phelan (1997). This value is higher than that fixed by van der Klaauw and Wolpin (2005) who fix it at 0.95, but lower than the rates estimated by French (2005), who find 0.985–

[13]There is a brief discussion about this in footnote 21 of Lee (2005) related to Lee's weighting strategy, but no information is provided about the calculation of standard errors.

[14]Running on 9 processors, estimation took about a week, with a sequence of simplex and gradient-based methods.

Table 5: Empirical moments

aggregate moment	#	units	weight
mean earnings by sector, age/cohort, education and experience	$2 \times (2 \times 13 + 14) \times 10 \times 5$	10^5NOK98	0.5
(mean earnings)^2 by sector, age/cohort, education and experience	$2 \times (2 \times 13 + 14) \times 10 \times 5$	10^{10} NOK^{2}98	0.25
choice distribution by age/cohort and education	$4 \times (2 \times 13 + 14) \times 10$		5
choice transitions by age/cohort	$4 \times 4 \times (2 \times 13 + 14)$		2
employment transitions by age/cohort, experience and pension point score	$2 \times 2 \times (2 \times 13 + 14) \times 5 \times 4$		5
total number of moments before restrictions	13440		

Note: There are some further restrictions: no empirical moments calculated with experience or education larger than age. Individuals only have 25 years to go to school, the 10 last years before legal retirement age sector change is restricted to zero. No-one can transfer from illegal states. In all, 10508 moments are given positive weight in estimation.

1.04. State-dependent transition probabilities for the value of home production turned out to be hard to identify, I have restricted the value of home production to be constant.

The two types are distinctly different. Type 1 has a fairly flat earnings-profile with small returns to experience and small levels of sector specific skills. Type two starts with a much lower base level of skills, but has the possibility of gaining large amounts of sector specific skills – these almost triples the skill-level. These large difference between the types, together with the fact that the type-shares for the 1941 cohort is very different from those of the 1955 and 1970 cohort may indicate that there is in fact room for more heterogeneity in the model, and that I should attempt to estimate model with three types.

The probability of gaining the sector specific skills are slightly lower in school (at 0.059, an expected length to gain skills of 17 years) than at work (at 0.066, or an expected length to gain skills of 15 years). But this difference is not large, and those who gain skills at school accumulate "general skills", in the sense that they are free to choose what sector to apply them to, whereas the skills gained at work are sector specific. There is quite a high probability of losing skills that are not used (at 0.42, or an expected length to loss of skills of 2.4 years). This creates a significant skill-based lock-in for type 2 individuals who have reached the high-skill level.

Most of the parameters that are not close to being so small as to have no economic impact on choices are estimated at a moderately comfortable level of precision. We should not be surprised that parameters that estimated to have little or no effect on choices or observed earnings (such as the correlation between the shock to public and private sector wages) are imprecisely estimated.

5 Policy experiment

The "Johnsen-commission", with a mandate to examine the National Insurance scheme has suggested several changes to the current scheme, to be implemented from 2010, starting with the 1951 cohort. Many of the suggested changes are much to fine-grained to be captured well by the stylised model in this paper, but some of the larger feature can be captured here: The commission proposes to abolish the 20-best-years averaging of the accumulated pension points, and a

Table 6: Structural estimates, common parameters

parameter	value
smoothing factor, ρ	8.24e-6
	(0.67e-6)
discount factor, δ	0.98
	(fixed)
technological progress, ξ	13.91e-3
	(1.14e-3)
utility of school, u_3	160e3
	(15e3)
cost of returning to school, u_{3r}	115e3
	(30e3)
prob. of losing non-used skill	0.416
	(0.068)
prob. of gaining skill by work	0.066
	(0.017)
prob. of gaining skill at school	0.059
	(0.011)
base productivity in home production	148e3
	(9e3)

Table 7: Structural estimates, type shares

Demographic group	Share of type 1
1941 cohort	0.698
	(0.046)
1955 cohort	0.341
	(0.047)
1970 cohort	0.279
	(0.053)

Table 8: Structural estimates, public and private sector skill equation

	public sector		private sector	
parameter	type 1	type 2	type 1	type 2
constant	125e3	83.7e3	115e3	75.3e3
	(9.3e3)	(8.1e3)	(13.2e3)	(12.9e3)
sector-specific skill	0.032	0.958	0.184	1.088
	(0.685)	(1.267)	(0.089)	(0.176)
own experience	1.73e-3		16.0e-3	
	(2.40e-3)		(3.24e-3)	
own experience squared	-0.10e-3		-0.40e-3	
	(0.04e-3)		(0.11e-3)	
experience in other sector	1.72e-3		0.10e-3	
	(1.96e-3)		(1.7e-3)	
σ_ε	0.269		0.019	
	(0.046)		(0.820)	
$\text{corr}(\varepsilon_1, \varepsilon_2)$	0.860			
	(31.9)			

replacement by a flat function of past income. It also proposes to replace the early retirement scheme (AFP) with a flexible choice of retirement age, where individuals are less favourably compensated when they retire early. One would think it would be of value to know to what extent this would affect retirement decisions. The analysis in Norges offentlige utredninger (2004) is based on analysis using "MOSART" (Fredriksen 1998). It is based on the assumption that future choices under the new system is a simple average of current conditional choices and the choice from before the introduction of AFP.[15] As of today, the agreement between the political parties (Arnstad, Nilsen, Steensnes, Stoltenberg and Vihovde 2005) does not indicate an abolishment of AFP (although adjustments may be made).

The details of the agreement among the political parties is hard to interpret in terms of the primitives of my model, e.g. it includes a lower rate of indexation, in which the average of wage growth and inflation is used for adjusting the value of benefits, and I have not attempted to model a change in the life expectancy as forecasted by the public commissions. As a first approximation,

[15]Personal communication with Dennis Fredriksen, Statistics Norway.

I model only a simple "core" reform. I model the removal of the "twenty-best-years" rule, and I adjust the National Insurance down slightly. I re-estimate $P_{\mathrm{pp}^f}(\mathrm{pp}^f(\theta_{t+1})|\alpha_t, w_{\alpha_t}(\theta_t), \mathrm{pp}^f(\theta_t), t, d)$, the transition function for the final pension point score. I use the same cut-offs as earlier, but with a flat average of past pension points instead of the "twenty-best-years" rule (allowing for more "down" movements on the fixed pension point grid). I also reduce the National Insurance benefits by a modest 5%. In Figure 7 one can see the effect of this on the out-of sample predictions for the 1970 birth cohort.

First, we can note that the overall employment levels seem rather low. However, we can note that the model does not allow for part-time workers (who are classified as primarily doing home production). The jaggedness of the labour supply profile is an artifact of the Markov transition matrix only being updated every fifth year. The overall pattern under the baseline is in line with the historical experience: A more or less flat participation in the public sector, with a decrease in the private sector over time. This decrease in the private sector is reasonable since there are weaker pension capital benefits in the private sector. There is, however, a peculiar influx into the public sector at the end of the lifecycle: These are people who are transferring to the public sector to be guaranteed AFP – early retirement benefits in the last year such a transfer is possible (by a restriction in the model). Since I have no firms in the model, I let private sector workers face a gamble on whether they could take AFP retirement, and it seems as if the value of AFP is to great for people to take that risk. This is not in itself alarming, since a number of other papers have shown that there are extremely strong incentives for people to take up the AFP if possible, and people who put any value on leisure will be predicted to retire early unless there are outside restrictions on the possibility of gaining a AFP-qualifying job.

The effect of the policy experiment is to reduce the labour supply of the middle aged group. Because of the relatively high rate of technological progress estimated, the retirement benefits seem to have early effects on labour supply. This is in line with what Rust (2001) has found. In my model it might be further suppressed by the fact that people realise that the most likely retirement option is with AFP from the public sector, and hence weights life-cycle contributions less. The policy experiment has a further effect at the end of the lifecycle. There is a marked increase in labour supply in the private sector among people in their mid-sixties. These people are the ones who did not win out in the AFP gamble – either they stayed on in the private sector, or they did not get the high public-sector wage draw they had hoped for, and which would have made them retire early with a high level of benefits.

6 Conluding remarks

I have estimated a basic two-sector life-cycle labour supply model with only two inital types of workers. However, as time goes on and individuals make different choices, a large amount of ex-post heterogeneity is created. The estimates indicate that there is a type of worker for which comparative advantage is generated by their accumulated life-history, and for which the endogenous costs of switching sectors can be large. The model extends the human capital literature that takes education as a deterministic determinant of wages into a more realistic environment where education is one of the factors (together with work experience) that goes into producing skills that are valued by workplaces – without guaranteeing a monetary payoff of education.

The policy experiment emphasises the fact that the labour supply effects of a National Insurance reform cannot be analysed entirely independent of the benefits in the public sector, and that the early retirement interacting with public sector benefits can have large effects on incentives. We cannot, of course, expect to see the extreme influx of workers to the public sector predicted by the model in real life: The model supports the conclusion of earlier studies that the incentives to retire using the AFP scheme are strong, and the fact that people do continue to work after qualifying for early retirement is hard to reconcile with a positive value on home production.

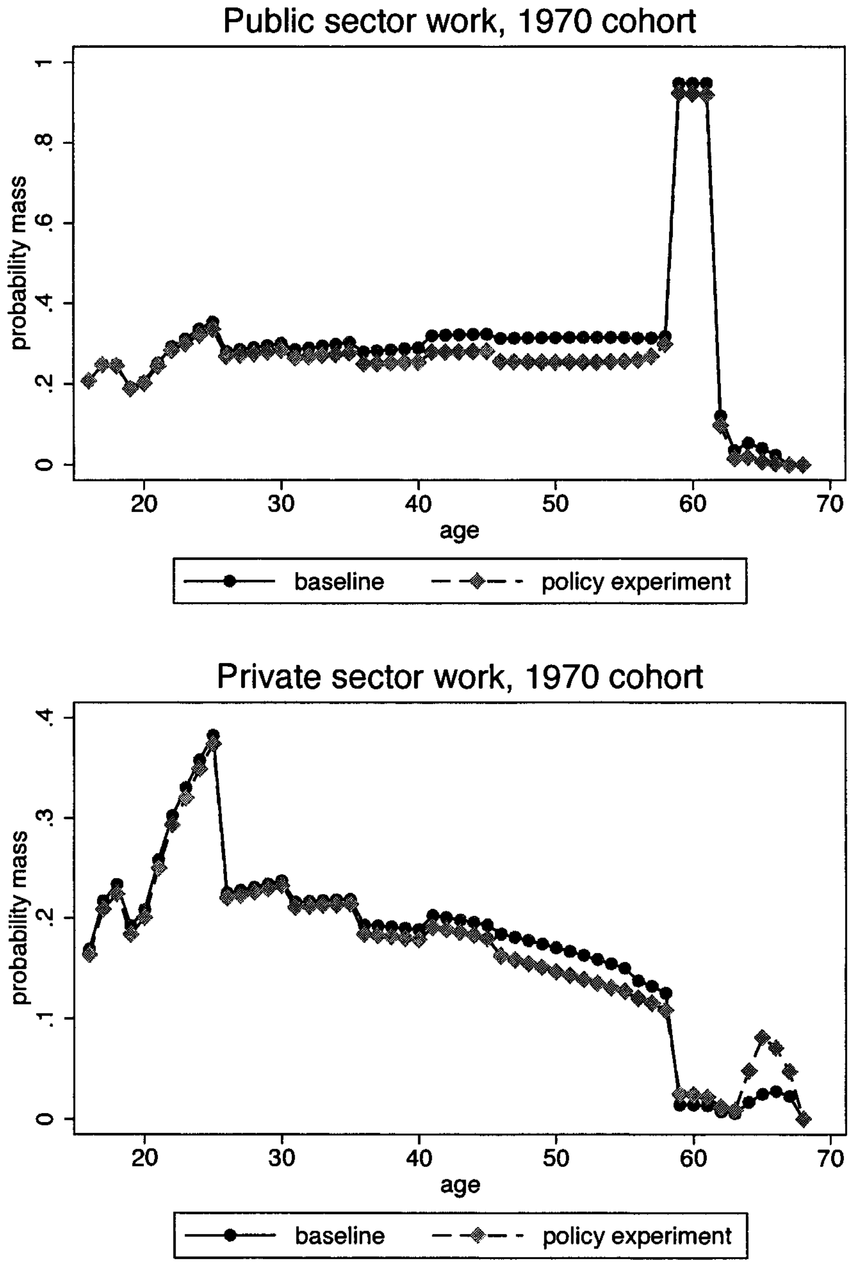


Figure 7: Policy experiment. A removal of the "twenty-best-years" rule and a 5% reduction in National Insurance benefits. The "baseline" is the current policy, under which the model is estimated.

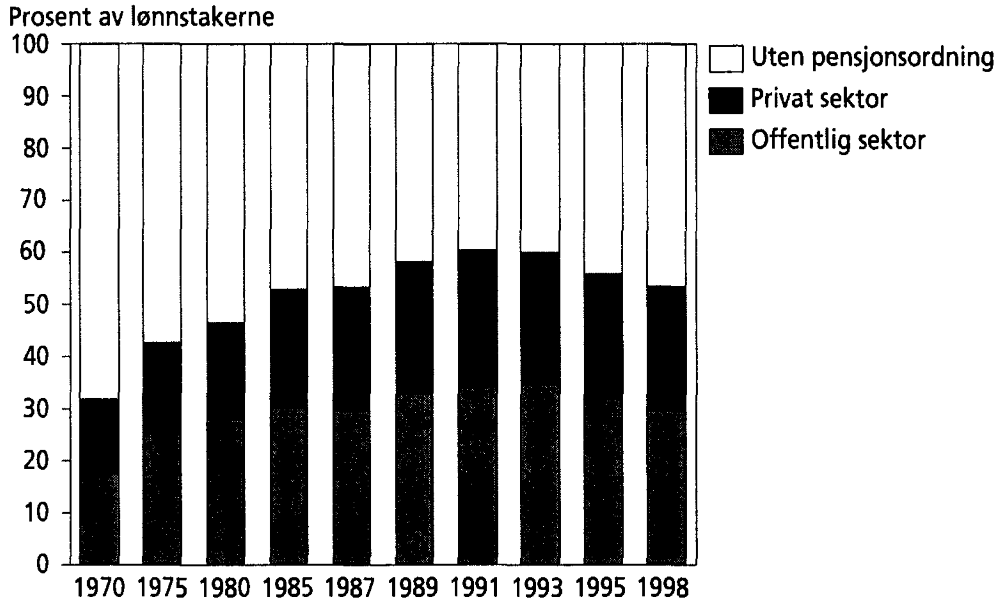


Figure 8: Share of workers covered by supplementary pension plans. Public sector at bottom, private sector in the middle and no coverage on top. The figure is taken from Pedersen (2000) and is based on Labour Force Survey data (AKU).

Appendix A Institutional setting: The Norwegian pension benefits

In this section I describe the Norwegian pension benefits at a level of some excess detail, to make it possible to evaluate the representativeness of the incentives that are embedded in my empirical model.

There are many different private sector schemes. The great majority are also of the defined benefit type, but employers typically pay one of the large insurance houses to provide actuarial annuities, and firms contribute based on the characteristics of individual workers. These terms are much like what individuals can obtain in individual retirement savings schemes. If a worker leaves a firm before retirement, the insurance firm providing the retirement issues a bond guaranteeing an annuity at retirement corresponding to the contributions of employers. Not everyone are covered by the private sector supplementary pension schemes, but if a corporation choose to provide a pension scheme, it must do so to all employees in order to expense contributions before taxes.

A.1 National insurance

The National Insurance pension benefits consists of a basic pension provided to all with more than three years of residence.[16] The legal retirement age is 67. There are three elements in the National Insurance old age pension. There is the **basic pension**, the **special supplement** and the **supplementary pension**. The basic pension and the special supplement together makes up the **minimum pension**, and is not dependent on past contributions to the national insurance system, but is available to all residents. If they have lived less then 40 years in Norway, the minimum pension is reduced accordingly. For national insurance and public sector pension benefits, all earnings (no capital income, no pension benefits) are calculated in units of **basic amount** ("grunnbeløp") G, the unit of account. Inflation is corrected for by adjustments of G, sometimes several times a year, but G is also meant to follow the average wage growth, although it has grown slower than wages historically, and only since the last revision in 1992 of the National Insurance in 1992 has there been any significant real growth in the value of a basic amount.

The basic pension is $1G$, except when the pensioner has a spouse who works and earn more than $2G$ or when the spouse is retired, in which case the basic pension is $0.75G$ for each.

For single persons or for people with non-retired spouses, the full special supplement is $0.7933G$. If the pensioner is providing for a spouse over the age of 60, the couple together gets $1.5866G$. For a couple where both spouses are retired, the supplement is also $0.7933G$ unless one them has a higher supplementary pension, in which case the special supplement to the spouse is $0.74G$.

The supplementary pension is based on earned incomes in the past. Earnings from a given year is first mapped into a number of **pension points**. Over time, there have been some changes in how pension points have been calculated from earnings, and Figure 9 shows how the progressiveness has changed over time.

The yearly pension points are summarised by a **final points score**, pp^f. This is the average of the 20 highest yearly pension points earned. If a pensioner has earned pension points in $x < 20$ years, it is the average of these x years. The final supplementary pension in a year t is then $G_t \cdot \text{pp}^f \cdot \sum_j \pi_j/40$, where the summation j is over the working years but limited to 40 years. Work experience before 1992 is awarded a replacement rate $\pi_t = 0.45$, while experience from 1992 and onward is given a replacement rate $\pi_t = 0.42$. The national insurance old age pension is then the basic pension plus the maximum of the special supplement and the supplementary pension.

[16] For a more detailed description, see *The Norwegian Social Insurance Scheme 2002*, `http://www.dep.no/sos/engelsk/publ/handbooks/044051-120003/index-dok000-b-n-a.html`.

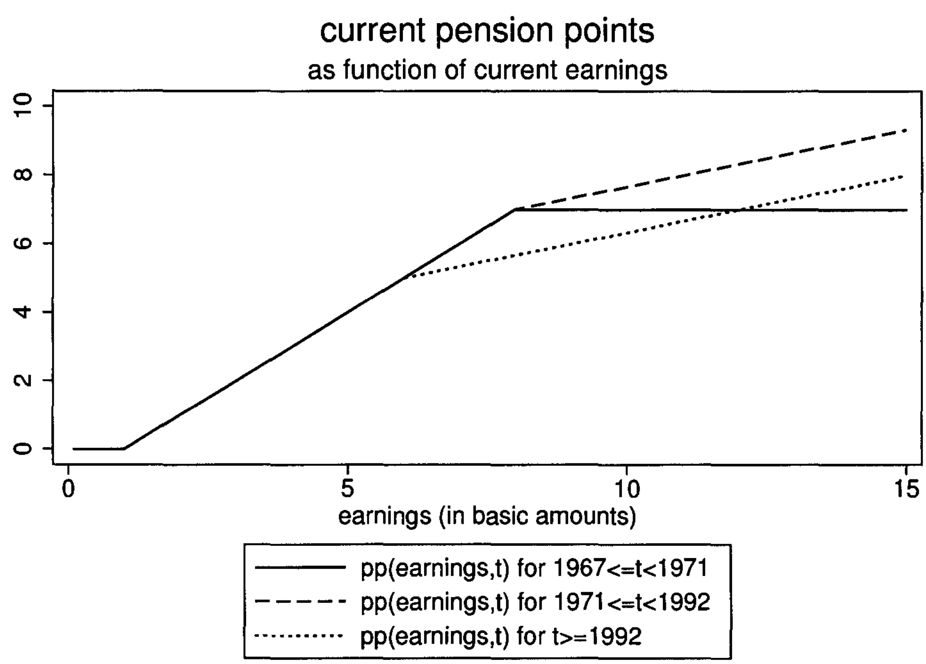


Figure 9: Pension points as functions of current earnings. Earnings in units of basic amount (G). In the year 20000, one basic amount was NOK 48400, approximately USD 7500.

A.2 adding public sector benefits

The public sector benefits are different. The system is complex, but the main characteristics for people who stay in the public sector is possible to explain. The various public sector insurance funds have agreements that the fund responsible for the individual when retirement takes place pays all benefits. For employees of the central administration, this system also applies for mobility between the Scandinavian countries.

There is a guaranteed benefit level of $0.66 \cdot \min(w^f, 12G_t)\lambda$, where w^f is the wage at the end of the career and $\lambda \in [0, 1]$ is the share of a "full career" worked in the public sector. For a "full career" in the public sector, only 30 years are needed – but if a person does not retire from the public sector, the "full career" career is calculated using between 30 and 40 years, depending on how long the actual career could have been if the retiree had stayed on in the public sector. If the ending wage is low in comparison with the wage over the rest of the career, adjustments will be made. For the years before the record wage, the record wage is applied. If a person reaches 30 years of experience at age 67, but there was a decrease in wage from 100 to 70 at age 60, the "final wage" will be calculated as $23/30 \cdot 100 + 7/30 \cdot 70 = 93$. But if the individual in fact has 37 years of experience in the public sector at retirement age and the record wage of 100 occurred at the end of a 30-year spell, the ending wage will be applied as 100.

The public sector benefits are only an addition to the national insurance, these two benefit levels are consolidated in ways that most Norwegians find somewhat cryptic.[17] An individual is guaranteed the maximum of national insurance benefits and public sector pensions, and the specifics is such that usually people will get more than this. The pension paid by the public sector pension fund is first reduced by $0.75 \cdot \lambda \cdot G_t$, corresponding vaguely to the basic pension. The supplementary pension from the national insurance is in principle deducted by a λ factor, unless parts of it was earned in the private sector. In this case, the part earned outside is not included in the deduction from the public sector. The calculation of the deduction is "simplified" by using the final wage to calculate a "fictitious" number of pension points, and the pensioner is allowed to keep an amount $(w^f - G^f) \cdot y^p/40 \cdot \pi \cdot G_t$ of the national insurance benefit.

To a first approximation, the pensioner gets the maximum of his guaranteed level from the public sector fund and what he is earned by the national insurance.

A.3 private sector benefits

Less than half of all those employed in the private sector are members of corporate sponsored pension plans, see Figure 8 taken from Pedersen (2000). In most cases, the firms buy these pension services from private insurance companies.

[17] It is regulated in *Lov om samordning av pensjons- og trygdeytelser*, 1957-07-06 no. 26, `http://www.lovdata.no/all/hl-19570706-026.html`.

While there are peculiarities with some of these programs, these programs are not consolidated with the national insurance. If an employer leaves a firm to work at another, the value of past pension contributions is calculated and a special retirement bond is issued. Individuals are allowed to continue paying into the programme on the same rate as those the corporation paid, but the next firm the individual applies to can also take on the previous pension contributions and respect those as if they were contributions into the new firm's programme. And lastly, the individual can hold on to the bond as such as cash in the stream of payments which will start at the official retirement age of 67.

Most private pension funds provide pension programmes which mimic the public funds (a fixed proportion of final wage), but at a corporation level. Also the guarantee is not absolute: One is in fact provided an annuity of the difference between the fixed percentage of the final wage and the national insurance level at the time of retirement, while the public programmes will compensate should the national insurance benefits decline in the future. If one should leave early in these programmes, ones future benefits are converted into a transferable promise of an annuity, but some corporations (e.g. Storebrand) advertises that a benefit to the corporation of these are more loyal employees. This has to do with how contributions and benefits were calculated in the past and there have in fact been clauses that firms can refuse to give older newly hired workers access to the corporation's pension programme. They have had their reasons for this: The contributions firms have made on behalf of their employees have traditionally been non-linear and concentrated in the latest years of a worker's career. In Figure 10 I have predicted typical age-profiles of contributions using the "annuity-method" (Clausen and Hersoug 2003).[18] This method has been the most popular one, since this method was required for tax exemption, and we see that it reaches 5% of earnings only as people are in their fifties.

A.4 early retirement

There are some local forms of early retirement possible. But in later years, a major policy innovation has been the AFP program. The "AFP" (Avtalefestet Pensjon) was introduced as an agreement between the parties in centralised wage-bargaining. Today, the programme allows retirement from the age of 62, but is not extended to everyone. It is extended to all of the public sector, and to the firms affected by private sector agreements between NHO and LO (the two main employer and employee organisations involved in wage bargaining). When it was introduced around 1990, it only allowed for retirement one year early, and has then been gradually extended. The latest extension (from 64 to 62 years) was in March 1998. Unfortunately, it not clear from our data which firms in the

[18]These are calculated by estimating a 4th degree polynomials of earnings on age on those of the 1937 cohort employed in private firms in 1986, then using this polynomial to calculate benefits.

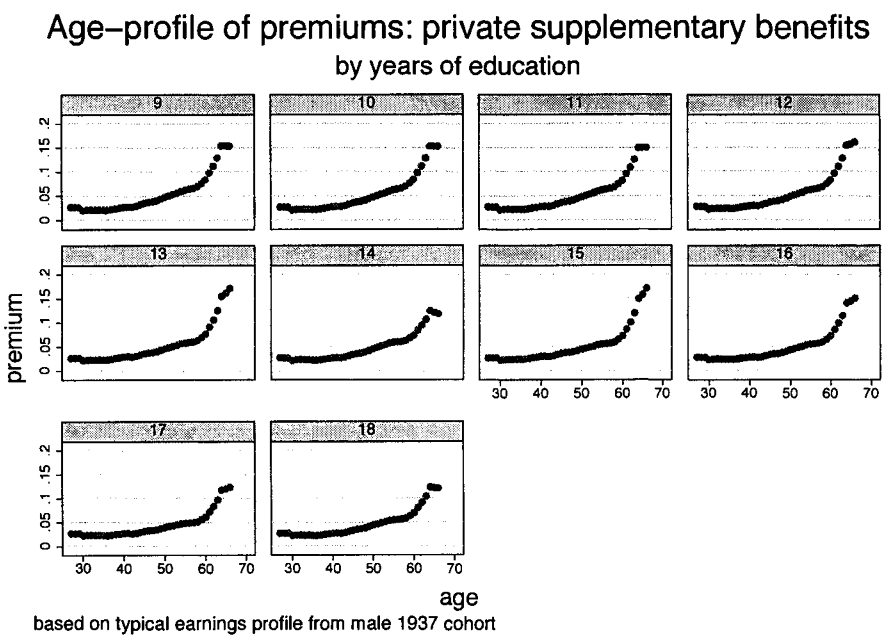


Figure 10: Typical private supplementary benefit premium profile

private sector is covered by the central wage-bargaining agreement. A further complication with this programme is that it depends on your history with the latest *firm* you are attached to: you need at least 3 years with the latest firm or 5 years with firms covered by the AFP agreement.

References

Arnstad, M., O. Nilsen, E. Steensnes, J. Stoltenberg and M. B. Vihovde (2005). "Pensjonsavtale mellom regjeringspartiene og Ap/Sp." Oslo, 19. mai.

Bratberg, E., T. H. Holmås and Ø. Thøgersen (2001). "Assessing the effects of early retirement programs." mimeo, Norwegian School of Economics and Business Administration.

Brinch, C., E. Hernaes and S. Strøm (2001). "Labour supply effects of an early retirement programme." CESifo Working Paper No. 463.

Clausen, J. R. and T. Hersoug (2003). *Bedriftens pensjonshåndbok.* Gyldendal.

Domeij, D. and P. Klein (2002). "Public pensions: To what extent do they account for Swedish wealth inequality." *Review of Economic Dynamics*, **5**(3), 503–534.

Ferrall, C. (2002). "Estimation and inference in social experiments." mimeo, Queen's University.

Ferrall, C. (2005). "Solving finite mixture models: Efficient computation in economics under serial and parallel execution." *Computational Economics*, **25**(4), 343–379.

Fredriksen, D. (1998). "Projections of population, education, labour supply and public pension benefits: Analyses with the dynamic microsimulation model MOSART." SØS 101, Statistisk sentralbyrå.

French, E. (2005). "The effects of health, wealth, and wages on labour supply and retirement behaviour." *Review of Economic Studies*, **72**(2), 395–427.

Hægeland, T., T. J. Klette and K. G. Salvanes (1999). "Declining returns to education in Norway? comparing estimates across cohorts, sectors and time." *Scandinavian Journal of Economics*, **101**(4), 555–576.

Heckman, J. and B. Singer (1984). "A method for minimizing the impact of distributional assumptions in econometric models for duration data." *Econometrica*, **52**(2), 271-320.

Hernæs, E., M. Sollie and S. Strøm (2000). "Early retirement and economic incentives." *Scandinavian Journal of Economics*, **102**(3), 481–502.

Jia, Z. (2005). "Retirement behavior of working couples in Norway. a dynamic programming approach." Discussion Paper 405/2005, Statistics Norway.

Keane, M. P. and K. I. Wolpin (1997). "The career decisions of young men." *Journal of Political Economy*, **105**(3), 473–522.

van der Klaauw, W. and K. I. Wolpin (2005). "Social security, pensions and the savings and retirement behavior of low income households." Penn Institute for Economic Research 05-020.

Lee, D. (2005). "An estimable dynamic general equilibrium model of work, schooling, and occupational choice." *International Economic Review*, **46**(1), 1–34.

Lee, D. and K. I. Wolpin (2004). "Intersectoral labor mobility and the growth of the service sector." forthcoming in *Econometrica*.

Midtsundstad, T. (2004). "Hvor mange har rett til AFP?" FAFO, rapport 728.

Møen, J., K. G. Salvanes and E. Ø. Sørensen (2003). "Documentation of the linked employer-employee dataset at Norwegian School of Economics." mimeo, Norwegian School of Economics.

Norges offentlige utredninger (2000). "Tjenestepensjonen etter skifte av arbeidsgiver og strukturendringer i offentlig sektor." Arbeids- og administrasjonsdepartementet, NOU 2000:28.

Norges offentlige utredninger (2004). "Modernisert folketrygd: Bærekraftig pensjon for fremtida." Finans- og Sosialdepartementet, NOU 2004:1.

Olsen, G. (1993). "Rutiner for ajourhold av det sentrale bedrifts- og foretaksregister." Statistisk Sentralbyrå, Notater, 93/13.

Pedersen, A. W. (2000). "Konkurranseutsetting og pensjon." FAFO rapport 338.

Raaum, O., K. G. Salvanes and E. Ø. Sørensen (2003). "The impact of a primary school reform on educational stratification: A Norwegian study of neighbour and school mate correlations." *Swedish Economic Policy Review*, **10**(2), 143–170.

Rønningen, D. and D. Fredriksen (2002). "Special tax rules for old-age pensioners." *Economic Survey from Statistics Norway*, **X**(2), 49–55.

Rust, J. (2001). "Modeling behaviorial responses to changes in social security: A life-cycle framework." mimeo, University of Maryland.

Rust, J. and C. Phelan (1997). "How social security and medicare affect retirement behavior in a world of incomplete markets." *Econometrica*, **65**(4), 781–831.

Statistics Norway (2001). *Statistical Yearbook*. Statistics Norway.

Statistisk sentralbyrå (1958). "Det norske skattesystemet 1958 - the Norwegian system of taxation 1958." Samfunnsøkonomiske studier 7.

Statistisk sentralbyrå (1968). "Det norske skattesystemet 1967 - the Norwegian system of taxation 1967." Samfunnsøkonomiske studier 20.

Statistisk sentralbyrå (1975). "Historisk oversikt over skattesatser m.v.:del I, årene til og med 1969." Arbeidsnotater IO 75/5.

Statistisk sentralbyrå (1988). "Skatter og overføringer til private: Historisk oversikt over satser mv. årene 1970-1988." Revidert utgave, rapporter 88/20.

Statistisk sentralbyrå (1994). "Skatter og overføringer til private: Historisk oversikt over satser mv. årene 1975-1994." Rapporter, 94/21.

Statistisk sentralbyrå (1995). "Historisk statistikk 1994 - historical statistics 1994." Norges offisielle statistikk, C188.

Statistisk sentralbyrå (2001). "Utdanningsstatistikk: Dokumentasjon 2000 av den individbaserte utdanningsstatistikken." Norges offisielle statistikk, C645.

Vassenden, E. (1995). "The register of the level of education: a product of the integration of other registers, census and surveys." Statistisk sentralbyrå, Notater 95/4.