The neighbourhood is not what it used to be: Has there been equalisation of opportunity across families and communities in Norway?[1]

Oddbjørn Raaum[2]   Kjell G. Salvanes[3]   Erik Ø. Sørensen[4]

21st December 2001

[1]We are grateful for comments from seminar participants at the Norwegian School of Economics, the University of Bergen, the University of Oslo and the Third Norwegian Meeting in Labour Economics at Jægtevolden. 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.

[2]Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, N–0349 Oslo, oddbjorn.raaum@frisch.uio.no.

[3]Department of Economics, Norwegian School of Economics and Business Administration, Helleveien 30, N–5045 Bergen, kjell.salvanes@nhh.no.

[4]Department of Economics, Norwegian School of Economics and Business Administration, Helleveien 30, N–5045 Bergen, erik.sorensen@nhh.no.

Abstract

Parents influence their children's adult outcomes through economic and genetic endowments, transmission of cultural values and social skills, and through choice of residential location. Using a variance decomposition framework which provides bounds on the effect of families and neighbourhoods, we find important effects of family characteristics as well as residential location on educational attainment and adult earnings in Norway. Families are more important than neighbourhoods as the correlations among siblings are significantly higher than among children growing up in the same local community. Sibling correlations are estimated to be a little lower than for the US, while correlations between neighbourhood children in Norway are found to be significantly weaker than in the US. Unlike previous studies, we also assess changes over time by studying children growing up around 1960 and 1970. While family effects are permanent over time, the impact of neighbourhoods is reduced by half in size from 1960 to 1970 and we link this result to several policy changes in the 1960s aimed at increasing equality of opportunity in Norway. Our results differ from previous US studies, suggesting that the role of families and neighbourhoods in explaining the degree of equality of opportunity and social mobility depends on labour market institutions and redistributive policies.

Key words: Education, Children, Neighbours, Siblings, Local Institutions, Peer-effects.
JEL classification: I21, J13, R23.

1 Introduction

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 routes: through investment in their children's education, through transmission of cultural values and social skills such as commitment to learning and social adaptability, and through genetic endowments. This early transmission of endowments and the investment of time and money in children are hypothesised to explain why some children achieve success as adults while others do not. Most studies show a strong degree of intergenerational transmission, since parental characteristics such as income and education are highly correlated with their children's outcomes along the same dimensions (see Solon (1999) for an overview).

Geographical location of the family is one specific aspect of parental behaviour. Researchers, social commentators and politicians often claim that the community where an individuals spend his/her childhood is likely to have effects on his/her adult socio-economic outcomes and well-being in general. Originally the interest in the effect of community background on adult socioeconomic outcomes stems from hypotheses about the effect of growing up in underclass neighbourhoods, but recently the scope of these analyses has broadened to include the effects of communities in general (see Solon, Page, and Duncan (2000), Page and Solon (2000), Ginther, Haveman, and Wolfe (2000) and Katz, Kling, and Liebman (2001) for recent papers along these lines). Neighbourhood effects is a label for a variety of different mechanisms. The attitudes and behaviour among peers with whom children and adolescents interact, the existence and enforcement of social norms, as well as local institutions like schools and childcare vary across neighbourhoods. 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 (dis)advantaged families into (dis)advantaged neighbourhoods. The variation in local institutions such as schools and nurseries may also increase when sorting intensifies, and this may 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 to pay a premium on house prices and rental prices in order to get access to better schools and neighbourhood for your children, shows that residential location is perceived as important by the public (Black 1999; Gibbons and Machin 2001). Jargowsky (1996) finds some evidence of increased segregation by income in US metropolitan areas between 1970 and 1990.[1,2] However, the evidence supporting causal effects of childhood neighbourhood characteristics on opportunities or behaviour affecting adult outcomes is far from compelling. In particular, as Manski (1995) argues, it is difficult to statistically identify causal neighbourhood effects. 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 of neighbourhood from the impact of peer group behaviour on individual outcomes via social interaction.[3] 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. For instance, Ginther et al. (2000) state in their reexamination of the finding of neighbourhood effects that: "The coefficients on the neighbourhood variables tend to fall in value and lose statistical significance as the specification of family variables becomes more complete." In addition to high correlation between family and neighbourhood characteristics due to sorting, several other problems also arise when using observable measures on residential location characteristics: Which characteristics are important, and what biases result from an incomplete set of characteristics.

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 persons who spent their childhoods in the same local community may reflect family influence (nature or nurture) rather than true neighbourhood effects. And vice versa, sibling resemblance can be due to neighbourhood effects, since most brothers and sisters grow up in the same community.

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 (see Solon et al. 2000; Page and Solon 2000; Solon 1999, for an overview). While the neighbour correlation represents an upper bound on the fraction of neighbourhood effects in the total variance, the sibling correlation measures the influence of all characteristics that siblings share.[4] The idea is simply that if aspects of the family and/or residential community during childhood and adolescence are important determinants of adult outcomes,

[1]Kremer (1997) on the other hand is more sceptical to 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]In addition to residential sorting, support is also found for increased marital or household sorting and sorting into schools, as well as increased sorting of workers into firms.

[3]See Katz et al. (2001) for an interesting attempt at controlling for sorting by using a social experiment design from the "Moving to Opportunity project" in Boston.

[4]The research strategy is explained in detail in Section 2, following Solon (1999).

there will be a strong correlation between siblings as compared to two arbitrarily chosen individuals. And if neighbourhood effects are important, those who become (un)successful adults would tend to have spent their childhoods along with neighbouring children who also end up (un)successful later on.

In the present paper we focus on the effects of family background and childhood location on adult educational attainment and earnings using data from Norway. Further, we contribute to the literature by assessing whether the effect of neighbourhood and family have changed over time. We expect institutional differences between societies, e.g. wage-setting institutions and governmental policies in general, to have an important influence on mechanisms determining social mobility and therefore also on equality of opportunity in a society. Studies from the United States seem to dominate the existing evidence on the effect of family characteristics on adult outcomes, and especially the effect of community background has almost exclusively been studied within the context of the US.[5] Using data from Norway with a different income distribution and institutional setting for wage determination and governmental social policy will add to the currently scarce evidence from Europe on neighbourhood effects.[6] In addition, using a data set on the population of the Norwegian citizens in 1960 and 1970, we are able to construct families and complete neighbourhoods both in 1960 and 1970, and to measure adult outcomes in 1990-1995.

This rich data set enables us to 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? 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 under different institutional settings. And furthermore, since we have data on families and neighbourhoods for two points in time, 1960 and 1970, we can compare family and neighbourhood effects over time and assess their stability by studying different birth cohorts.

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, implemented to increase equality of

[5]Using UK data, Ermisch and Francesconi (2001) and Dearden, Machin, and Reed (1997) are exceptions to the rule of American studies of family background on educational attainment and earnings. Gibbons (2001a,b) finds support for neighbourhood effects on educational attainment above family effects using UK data for 1970.

[6]See Soltow (1965) for the development of inequality in Norway historically, and for recent results see Aaberge, Björklund, Jäntti, Pedersen, Smith, and Wennemo (2000). For discussions of different institutional settings for wage determination and other institutions see Kahn (1998) and Hægeland, Klette, and Salvanes (1999).

opportunity via education reforms and other policy reforms. For instance, in the 1960s a comprehensive school reform took place in Norway extending mandatory schooling from 7 to 9 years (see Aakvik, Salvanes, and Vaage (2001) for an analysis of this reform). Additional social reforms that took place in the 1960s which may have influenced the effect of neighbourhoods include increased access to student grants, as well as strong measures to redistribute resources across municipalities, thus providing more equal distribution in local institutions, such as access to primary schooling and health care services.

The paper 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 3 comprises a data description, including how neighbourhoods and families are constructed, as well as a detailed description of the estimation procedures. The results are presented and discussed in Section 4, and Section 5 contains concluding remarks.

2 Statistical Model

In light of the problems using conventional regression analysis to identify various neighbourhood effects and previous failed attempts to disentangle them from family effects and selection on unobservables into local communities, we follow an approach suggested in several papers by Solon et al. (2000).[7] A simple statistical framework based on Solon et al. (2000) will be presented in this section to show how family and neighbourhood play a role in determining educational attainment and adult permanent income. First, we show how family and neighbourhood effects may be partially disentangled and how the upper bound of the neighbourhood effect may be restricted. Second, we describe how sibling and neighbour correlations can be combined to provide a lower bound on the variance of the family effects.

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 suppose that we can decompose additively the role of neighbourhood, family and individual-specific factors on the outcome as

$$y_{cfs} = \beta' Z_c + \alpha' X_{cf} + \epsilon_{cfs}, \tag{1}$$

where X_{cf} is a vector of all family characteristics, measured and unmeasured, that influence permanent earnings or years of education, Z_c contains all the neighbourhood characteristics, and ϵ_{cfs} representis unrelated individual factors orthogonal to both family and neighbourhood effects. Since we think of Z_c and X_{cf} as latent vectors that include all relevant variables, it is not a restriction to let ϵ_{cfs}

[7]Notable are also recent developments using IV estimators or quasi experiments to separate neighbourhood effects from family effects within a regression framework. See Gibbons (2001a), and Katz et al. (2001).

be orthogonal to the explanatory variables. We expect the family background effects, $\alpha' X_{cf}$, and the neighbourhood effects, $\beta' Z_c$, 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.

Now, if all the family and neighbourhood factors were observable and could be measured with accuracy, the strategy would be to estimate (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, we will show that it is possible to bound the variance of neighbourhood effects by measuring sibling and neighbour covariances in y_{cfs}. The idea is that if family and residential community are important in explaining adult status, there will be a strong correlation between siblings in their adult outcomes, as compared to two arbitrarily chosen individuals. And if the neighbourhood where the child grew up is important, it will show up in a strong correlation between neighbouring children's adult socioeconomic outcomes.

The total variance of the socioeconomic outcome, y_{cfs}, of all the individuals in the sample can be decomposed as:

$$\operatorname{var}(y_{cfs}) = \operatorname{var}(\beta' Z_c) + \operatorname{var}(\alpha' X_{cf}) + 2\operatorname{cov}(\alpha' X_{cf}, \beta' Z_c) + \operatorname{var}(\epsilon_{cfs}). \tag{2}$$

Similarly, the covariance between sibling s and s' is

$$\operatorname{cov}(y_{cfs}, y_{cfs'}) = \operatorname{var}(\beta' Z_c) + \operatorname{var}(\alpha' X_{cf}) + 2\operatorname{cov}(\alpha' X_{cf}, \beta' Z_c). \tag{3}$$

The sibling correlation, $\operatorname{cov}(y_{cfs}, y_{cfs'})/\operatorname{var}(y_{cfs})$, measures the proportion of the total variance in the outcome under consideration — in our case long term earnings or educational attainment — due to factors shared by siblings. From (3) we see that siblings have correlated adult outcomes because they shared neighbourhood characteristics, and family characteristics, second term. The sorting of families into neighbourhoods is reflected in the third term. The sibling covariance then captures all measured and non-measured factors shared by siblings that may have an impact on later performance, such as family status, genetic traits shared by siblings, family composition, as well as neighbourhood effects and sorting. Thus, the sibling correlation is an upper bound for family effects since the covariance also includes neighbourhood characteristics such as the quality and availability of community institutions and the characteristics of the other adults and children living in the neighbourhood. Although we cannot disentangle the family and neighbourhood characteristics in (3), more information can be gained by the neighbour covariance from different families in the same neighbourhood,

$$\operatorname{cov}(y_{cfs}, y_{cf's'}) = \operatorname{var}(\beta' Z_c) + \operatorname{cov}(\alpha' X_{cf}, \alpha' X_{cf'}) + 2\operatorname{cov}(\alpha' X_{cf}, \beta' Z_c). \tag{4}$$

In (4), we notice that the neighbour covariance consists of more than the variance in neighbourhood characteristics given in the first term and it should therefore be viewed as an upper bound on the neighbourhood influence on the covariance in y_{cfs} between neighbours. The second and third terms are both expected to be positive, leading to an upward bias. The second term represents sorting of similar families into the same areas, since neighbouring children share similar family characteristics. Similarly, the third term also represents sorting, in that it denotes sorting of (dis)advantaged families into (dis)advantaged neighbourhoods. We see that positive sorting, $\text{cov}(\alpha' X_{cf}, \alpha' X_{cf'}) \geq 0$ and $\text{cov}(\alpha' X_{cf}, \beta' Z_c) \geq 0$, implies that $\text{var}(\beta' Z_c) \leq \text{cov}(y_{cfs}, y_{cf's'})$.

The neighbourhood correlation, $\text{cov}(y_{cfs}, y_{cf's'}) / \text{var}(y_{cfs})$, 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. In particular, we estimate the joint effects of the variation in neighbourhood characteristics (Z's) and the causal effect they have on adult outcomes (β's). Consequently, if neighbourhood correlations change over time, or differ across countries, this may reflect another dispersion of relevant neighbourhood characteristics or differences in their effects. In the case of adult earnings, the effects are again a product of neighbourhood effects on human capital (skills) accumulation and the market prices of human capital. Comparing Scandinavian countries and the United States, for example, we may speculate that all three components are different.

Access to neighbourhood identifiers and family characteristics in the same data enables us to tighten the upper bound on the neighbourhood effect and also to establish a lower bound on the family effects. First, it follows from (4) that the upper bound on the neighbourhood effects can be made tighter by introducing observable family characteristics shared by the neighbours, and by subtracting that as an observable part of the second term of (4). Following Solon et al. (2000) and Altonji (1988), we estimate the part of $\alpha' X_{cf}$ related to observable family characteristics such as mother's and father's education, income, and the number of children in the family, as well as a dummy for divorced or separated families. Let $\tilde{X}_{cf}$ denote the observable subset of family characteristics with associated parameters $\widehat{a}$ estimated within neighbourhoods. We can then deduct the sorting component arising from the fact that similar families tend to cluster in neighbourhoods,

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

We expect the contribution from unobserved family characteristics to be positively correlated with our observed measure. Since we measure our observable characteristics with error, and we cannot expect to fully proxy unobservables with observables, it seems reasonable to suppose that $\text{cov}(\widehat{a}' \tilde{X}_{cf}, \widehat{a}' \tilde{X}_{cf'}) \leq \text{cov}(\alpha' X_{cf}, \alpha' X_{cf'})$. It then follows that

$$\text{var}(\beta' Z_c) \leq \widehat{\text{cov}}_{adj.}(y_{cfs}, y_{cf's'}) \leq \widehat{\text{cov}}(y_{cfs}, y_{cf's'}). \tag{6}$$

Even if all the relevant family characteristics were included in $\tilde{X}_{cf}$ and the associated parameter estimates were unbiased, the sorting of (dis)advantaged families into (dis)advantaged neighbourhoods would imply that the adjusted correlation represents an upper bound.

Second, the part of the sibling correlation arising from common neighbourhoods can be partialled out using the neighbour correlation. More precisely, by subtracting the adjusted neighbour correlation from the sibling correlation, what is left represents a lower bound on the variance of the family effect. Since the adjusted neighbour correlation is an upper bound on the neighbourhood effects, subtracting yields a lower bound. Combining (3) and (4) gives

$$\text{var}(\alpha' X_{cf}) = \text{cov}(y_{cfs}, y_{cfs'}) - \text{cov}(y_{cfs}, y_{cf's'}) + \text{cov}(\alpha' X_{cf}, \alpha' X_{cf'}),$$

and we can use the lower bound on neighbourhood effects (5) to identify the lower bound

$$\text{var}(\alpha' X_{cf}) \geq \text{cov}(y_{cfs}, y_{cfs'}) - \widehat{\text{cov}}_{adj.}(y_{cfs}, y_{cf's'}). \tag{7}$$

3 Data and estimation

Statistics Norway has provided the two sets of data we use. On one hand, we have a database of linked administrative data, which covers all Norwegians aged 16–74 with any earnings or registered unemployment in any of the years 1986–95.[8] On the other hand, we have extracts from the censuses in 1960, 1970 and 1980. In principle, the census-files provide information about the biological parents of the individuals in our administrative register data. The linking of administrative to census data is not perfect, as Table 1 shows, but for the subset of individuals we consider in this paper, more than 90% can be linked across these datasets. Statistics Norway consider that the main reason they are unable to link all individuals is that information about parents is lacking for many individuals in the administrative data. This is most pronounced for immigrants and for those who left home before 1960.

[Table 1 about here.]

The administrative data provide information about taxable income (excluding capital gains), educational attainment, labour market status and a limited set of demographical variables. From the census files we have information about the education of biological parents, occupation, municipality and census tract they lived in and their age. For the 1970 census we also have their income. Vassenden (1987) documents the census data.

[8] Comparing with national population statistics, it is clear that this database cover 99% of all males, and about 97% of all females in the cohorts we are examining.

3.1 Neighbourhoods

We use mother's census tract as identifier of neighbourhood. Where mother and father live apart, a very small number of children would live with their father in the years we consider, and consequently we expect only a small number to be mis-classified. 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 differ 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 1960's. Uniform guidelines for urban municipalities were not imposed, but the cities with large tracts in the 1960 census have mostly increased the number of tracts for the 1970 census. The tracts in Oslo, the capital of Norway, had an average of 4903 inhabitants in the 1960 census; this was reduced to 1091 in 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. In 1960, 6127 tracts had a population of fewer than 500 individuals. This number grew to 6809 in 1970.[9]

We observe the neighbourhood children live in 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. 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. 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. Because the tracts are not directly comparable across the two censuses, attempts were made to provide aggregations of tracts that are comparable.

Langen (1975, appendix D) provides a catalogue of 5298 such comparable units. In many circumstances there were no changes made to tract boundaries, and the "aggregation" consists of a single tract. But some tract aggregations are

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

very large (e.g. Oslo).[10] In order to examine how stable neighbourhood effects are, we will consider a subsample of these comparable units. We restrict the sample to units with fewer than 4000 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 0–5 in the 1960 census, who we can expect to live with their parents at the time of the 1970 census.

The Norwegian tracts were small by the international standards of the day. Sweden had 2568 "parishes" in 1971, with an average of 3145 individuals, Denmark had about 5000 primary units in 1970, with an average of 990 individuals. Great Britain had "enumeration districts" of about 750–1000 individuals, in the 1961 census (Langen 1975, p. 5–6). The US Bureau of the Census requires that the average population of all census tracts in a county be about 4000 people, and there were 62,276 tracts and Block Numbering Areas in the US 1990 census (Bureau of the Census 1994, p. 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 that can be identified from the PSID data, but smaller than the census tracts mostly used to assess neighbourhood effects using US data.[11]

3.2 Identifying families

The data we have from the censuses and the administrative records do not come with a family identifier. But as our census files refer to information about the *parents* of the individual we have identified in the administrative files, we can utilise this information. Sorting by parental characteristics, we can identify siblings as individuals with identical parents. We have census data from the 1960, 1970 and 1980 censuses, with information about census tracts for 1960 and 1970.

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

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

The characteristics we have include age, marital status, country of birth, education and occupation, and for 1970 and 1980 we also have income (in increments of NOK 100). Within census tracts, we consider the chances of two individuals accidentally having parents that are identical among all these dimensions to be quite small, and we conjecture that there is only modest mis-classification, indeed, using the 1960 census alone makes only minor difference to the family size distribution. Our largest constructed family consists of 16 children, while there are mothers with more children in the administrative data.

To evaluate our claim that this procedure is reasonable, we have compared the size distribution of the families we can construct this way with information from the administrative records about the total number of children born of women of different age groups. In Table 2 we see that if anything, our constructed families are slightly smaller than they should be, and that this is more pronounced for the older mothers. This reflects the fact that we are not able to link all individuals from the administrative data to the census files.

[Table 2 about here.]

3.3 Outcome variables and observed family background

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.

The measure we have of the educational attainment of the *parents* of individuals from our main sample is different. We have the original codes used in the 1960 and 1970 censuses. While we in principle could use dummy variables to correct for parental education, this would identify the coefficients from those neighbourhoods with two or more parents with the same educational background. For many educations, this is a tiny fraction of the census tracts. We would also like to compare educational attainment of parents across neighbourhoods. For these reasons, we have transformed the categorical education codes into years of education, using a two-step procedure. We map the codes from 1960 to 1970 codes using repeated observations of the same individuals, and translate into years using the oldest observations in the register of education at three-digit levels. We form the Cartesian product of two sets of educational codes in the 1960 census (basic and higher education). For each element in this set, we pick the mode of the three-digit 1970 codes for the individuals we observe both in the 1960 and 1970 census. This mapping is done separately for men and women. At three-digit level, the 1970 codes correspond to the codes in the register of education. We use three-digit codes for the 75 oldest completion dates of our individuals in the administrative data set to translate into years of education, and pick the median among these 75 as our best imputation of how many years the parents of our main sample went to school.

Our measure of adult earnings, or rather of the logarithm of adult earnings, use the register data which is based on tax returns. We use the 1990–95 observations of a category of taxable earnings that includes wages, income from self-employment, unemployment benefits and sick-leave payments, but excludes capital income, social assistance, pensions and other transfers. 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. We calculate the mean of the logarithm of these observations for each individual.

[Table 3 about here.]

[Table 4 about here.]

Table 3 provides summary statistics of our sample compared to the full population from the administrative data in 1995. We conclude that our samples are similar to the full population. Notice that our samples are only fractions of the full population in the relevant age group. In order to include a family in our sample, the family must include at least two children aged 5–15 at the time of the census. For the single-sex samples, we restrict ourselves to the families with at least two brothers or two sisters. This restricts the samples considerably, and it implies that the "all siblings" sample is not a weighted average of the single-sex samples. In addition to this, we restrict ourselves to neighbourhoods with at least two of these families. A number of small tracts do not satisfy this criterion. The fact that linking of administrative to census data is not perfect is of lesser importance. 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 (2000) use 443 individuals from 120 clusters to examine male earnings. We note that there is a limited increase in the average years of education from the older to the younger cohorts, and that 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 is measured at two different stages of the life-cycle. Table 4 provides summary statistics on parental characteristics.

3.4 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 (8) as an estimator of the covariance between siblings of a variable y with $\mathrm{E}(y) = 0$,

$$\sum_{c=1}^{C} W_c \left\{ \sum_{f=1}^{F_c} W_{cf} \left[\sum_{s \neq s'} \frac{y_{cfs} y_{cfs'}}{S_{cf}(S_{cf}-1)/2} \right] \Big/ \sum_{f=1}^{F_c} W_{cf} \right\} \Big/ \sum_{c=1}^{C} W_c. \tag{8}$$

Here c denotes cluster, 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 cluster c. The four weighting schemes used by Solon et al. are provided in Table 5. In practice, we have found the differences among estimates with different weighting schemes to be negligible. All estimates in this paper are with the fourth weighting scheme in Table 5, which means that all sibling-pairs and neighbour-pairs received equal weight regardless of whether they came from large or small families and neighbourhoods. To centre the observation around zero, we follow Solon et al. in first regressing the variable in question on age, and for mixed samples, the sex of individuals.

[Table 5 about here.]

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 (8),

$$\sum_{c=1}^{C} W_c \left\{ \sum_{f \neq f'} W_{cff'} \left[\sum_{s=1}^{S_{cf}} \sum_{s'=1}^{S_{cf'}} \frac{y_{cfs} y_{cf's'}}{S_{cf} S_{cf'}} \right] \Big/ \sum_{f \neq f'} W_{cff'} \right\} \Big/ \sum_{c=1}^{C} W_c. \tag{9}$$

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 the regular bootstrap, drawing neighbourhoods or families as clusters depending on which estimator we replicate. Our estimates of standard errors are based on 100 bootstrap replications. All the sample reductions and initial corrections are bootstrapped.

In order to adjust for the effect of observable family characteristics, we follow Solon et al. (2000) and regress y_{cfs} on a vector of observed characteristics $\tilde{X}_{cf}$ 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 observable family characteristics.

4 Results

We first present the unadjusted sibling and neighbour correlations and then we try to examine how much of these effects can be explained by the fact that families in a neighbourhood tend to be similar. We will briefly discuss how simple regional differences and neighbourhood misclassification affect our estimated neighbour correlations. We close this section with a discussion of policy changes during the 1950s, '60s and '70s that are candidates for explaining our results.

4.1 Sibling and neighbour correlations in education and earnings

The sibling and neighbour correlations are shown in Table 6. The brother correlation in years of schooling is around 0.43 and somewhat higher, 0.47, for sisters. These figures are surprisingly similar to what is found for other countries. According to Solon (1999), sibling correlations in years of education in the United States are a little higher than 0.5. Table 6 also reveals that the gender-specific education correlation is stable across cohorts. Thus, the overall impact of family and neighbourhood characteristics shared by siblings seems to be constant over time. The neighbour correlations in education are much lower, but clearly positive. While the male correlation is higher among those born 1945-55, there is no gender difference in the younger cohorts. While Solon et al. (2000) report an unadjusted neighbour correlation of around 0.2 in the United States, neighbourhoods seem to be less important in Norway.[12] 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 0.121 to 0.061 for males and from 0.109 to 0.062 for females.

[Table 6 about here.]

Correlations are considerably lower when we look at earnings. The sibling correlations in the range 0.16-0.21 are similar to figures found in previous studies. Björklund, Eriksson, Jäntti, Raaum, and Österbacka (2001) find that brother correlations in Scandinavia 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. Since female earnings are strongly influenced by labour supply decisions, partly caused by marital status and children, sisters are likely to show less resemblances than brothers.[13] The brother correlation drops significantly from the oldest to the

[12] Note, however, that the estimates of Solon et al. (2000) come from a much smaller sample, with standard errors around 0.05.

[13] Bound, Griliches, and Hall (1986) finds that sisters are *more* similar than brothers, but they examine residuals from wage equations.

youngest cohort, from 0.211 to 0.175, while the earnings correlation for sisters is constant over time, 0.156 and 0.157.

The neighbour correlations in adult earnings are clearly positive. Earnings correlations of neighbouring boys are higher than for girls who spend their childhood in the same local community. The higher male correlation is found in both cohorts. For the older birth cohorts, correlation in adult earnings among neighbouring boys is 0.068 and 0.033 for girls. As for education, the correlation is reduced by approximately one half from the 1945-55 to the 1955-65 cohorts.

To summarise, the raw 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.

4.2 Adjusting for family background

The neighbour correlations in Table 6 are clearly upwardly biased measures of the true influence neighbourhoods have on individuals, due to sorting of families into communities, see discussion leading to (6). 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 7 we report neighbour correlations in years of schooling where we subtract the covariance in effects of observed family characteristics.[14]

[Table 7 about here.]

When we partial out the effect of parental education, reported in column E, the correlation drops considerably. Family structure characteristics are less important, since the numbers in column D are very similar to the unadjusted figures. When combining the two sets of family background variables available for both cohorts, column $E+D$, the correlation drops considerably, from 0.121 to 0.043 for males and from 0.109 to 0.017 for females in the 1945-55 cohorts. The impact of the family background adjustment is similar for the 1955-65 cohorts. The gender difference found in the older cohort reappears among the younger, but the neighbour correlations for males and females have become more similar over time. Parental income information is only available for the younger cohort, but this adds little. The main results for educational attainment are two. First, partialling out the similarity between neighbouring children in the effects of standard family characteristics drives the neighbour correlation down by about two-thirds for men and even more for women. Second, the difference between cohorts remains.

[14] To prevent family background coefficients from being affected by neighbourhood effects, these are estimated within neighbourhoods.

[Table 8 about here.]

In Table 8 we report family background adjusted neighbour correlations in adult earnings. After partialling out the effects of parental education, column E, we find that the neighbour correlations fall in both cohorts and for both men and women. As for education, family structure is less important, see column D. Including parental income among the family controls has no substantial influence on the correlations.

Several important conclusions can be drawn from Table 8. First, observable family sorting into neighbourhoods does not seem to explain all the resemblance in adult earnings among persons who grew up in the same neighbourhood. Second, neighbourhoods have become less important as determinants of adult labour market success. This is consistent with, and presumably partly explained by, the declining neighbourhood effects 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, see columns $E + D$ in Table 7 and Table 8.

One might ask what correlation terms mean in terms of absolute size of neighbourhood effects. In Table 9 we have calculated the implied upper bound on the standard deviation of neighbourhood effects, using the correlations from Table 7, Table 8 and the corresponding measure of y. We see that although the correlations may seem low, the variance is large enough that the bounds on neighbourhood effects are in no way negligible.

[Table 9 about here.]

Previous studies of sibling correlations do not disentangle family from neighbourhood effects. The only exception, to our knowledge, is Page and Solon (2000). From Table 6, we can conclude that family background, whether it is nature or nurture, is by far more important than neighbourhoods. By combining estimates of the two correlations and the covariance in (observed) family characteristics within neighbourhoods, we obtain a lower bound on the overall family effects using equation (7). Our bounds are based on the same specifications, the $E + D$ columns, in both cohorts.

For educational attainment, the lower bound on family effects in the 1945-55 cohorts is 0.389 for males and 0.460 for females. In the younger cohorts, we find that at least 0.398 of the male variance in schooling is explained by factors, other than childhood location, which siblings share. The family effect remains stronger for females, with a lower bound at 0.462. Comparing the two cohorts, we find that the lower bound on family background effects, i.e. net of neighbourhood effects, is constant.

The neighbourhood-effect-adjusted brother correlation in earnings falls over time, from 0.157 among those born 1945-55 to 0.147 for those born ten years later. Consequently, a substantial part of the drop in brother correlation over time, from 0.211 to 0.175 in Table 6, is accounted for by the weaker neighbourhood effects in the younger cohorts. For females, the lower bound on family effects is also fairly stable: 0.136 in the older cohorts and 0.144 in the younger. This suggests that the higher (unadjusted) brother correlation reported in Table 6 for the youngest cohort, compared to sisters, is because neighbourhoods have a stronger effect on male than on female earnings. While sister resemblance is higher, compared to brothers, when we look at educational attainment, we find that the lower bound on family effects on earnings in the 1955-65 cohort is basically the same for both sexes.

4.3 Interpreting regional differences

Neighbourhood correlations in adult earnings may exist for numerous reasons. As pointed out by Griliches (1979, p. S38), in the case of brothers "...correlation arises from many sources: ...; the likelihood, even in adulthood, of closer location in space and hence exposure to similar regional price differentials and common business cycle effects; and more." Following up on Griliches (1979), several authors note that the measured neighbourhood effects can be due to the fact that there are different labour markets in different regions. It is well-documented that workers in urban areas in the US are paid a premium for living and working there.[15]

The explanations are complex, but the main point here is that the regional differences in industry structure and labour market characteristics leading to a difference in educational choices and long term earnings can be picked up as neighbour correlations in adult outcomes. If resemblance in adult earnings among neighbouring children simply reflects geographical location preferences heavily influenced by where they grew up, it has nothing to do with what people think of as neighbourhood effects, i.e. peer influences, social norm enforcement by local communities or quality of neighbourhood institutions like schools or nurseries/day care facilities.

Disentangling the potential impact of exogenous intergenerational location preferences from socially important neighbourhood effects is difficult. First, location can be causally influenced by the neighbourhood effects. To illustrate, assume that children from disadvantaged neighbourhoods rarely end up as adults in regions with high-paying jobs, partly because they had a bad start in life. In

[15]Using our measure of earnings in a Mincer earnings equation including schooling, experience, experience squared, sex and regional dummies, the largest difference is between the counties Akershus (high earnings) and Aust-Agder, a 21% difference.

this case, any adjustment for adult location will disguise some of the neighbourhood effects. Second, adult location may reflect unobserved characteristics of the individual, such as market luck or failure, or of the family. Consequently, average differences in adult outcomes may reflect unobserved heterogeneity and not "real" regional wage or cost of living differences. Third, identification is difficult and "exogenous" mobility is needed to separate the neighbourhood and regional effects from each other. Assuming that no children leave their childhood neighbourhoods when they become adults, i.e., that those who leave to go to school elsewhere return when they have completed schools or universities, there are no data to distinguish effects of childhood and adult location. Therefore, adult location adjustments must be based on the labour market success, or failure, of those who leave their childhood neighbourhood.

We acknowledge the problems associated with location preferences and perform a simple check where neighbour correlations are estimated within counties. If correlations in adult outcomes were completely driven by location preferences, within county estimates of the neighbour correlations should be zero. In Table 10 we estimate sibling and neighbour correlations based on age and childhood county adjusted adult outcomes. In the first step, schooling and earnings, respectively, are adjusted for year of birth and childhood regional dummies (20 counties). In this way we adjust for all effects of adult location that are explained by the region in which the person spent her childhood. If location preferences explained the correlations, we would expect the neighbour correlations to vanish. When we estimate within counties, we subtract the average neighbourhood effect in each region, i.e. all variation in neighbourhoods across counties is ignored. This will tend to bias the correlation downwards. By comparing Tables 6 and 10, we find that childhood region does not affect the sibling correlations in any significant way. The brother and sister correlations in adult outcomes are not explained by the fact that siblings spent their childhoods in the same region.

[Table 10 about here.]

The neighbourhood effects on educational attainment, however, drop significantly. Considering schooling years of the oldest cohorts first, the neighbour correlations are reduced by approximately 25% when we estimate within regions. Adjusting for childhood counties has no significant impact on the younger cohorts, although the within-county estimates are lower for both men and women, see Table 6 and Table 10.

Earnings correlation estimates are even more affected by the childhood-region adjustment, as expected. Regional wage differentials will partly be accounted for since adult location is strongly affected by childhood region. Comparing the lower panels of row one and four in Tables 6 and 10, we see that the neighbourhood earnings correlations drop by about 50%. Compared to Page and Solon (2000), where childhood urbanicity explains the same proportion of the adult earnings variance as neighbourhoods do, we find that childhood region plays a less important role in Norway. To illustrate, childhood regional dummies explain only 1.6% of the overall variation in earnings in the 1945-55 cohort. The within-county correlations in Table 10 suffer from the bias due to family sorting. In Table 11 we partial out the effects of observed family characteristics. This adjustment has minor effects for males, but the female neighbour correlations in earnings end up very close to zero. While no significant difference between the cohorts is found for females, the lower effects of neighbourhoods in the younger cohort remain for males. 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.

[Table 11 about here.]

4.4 Has the impact of neighbourhoods really declined?

We have found that neighbour correlations are lower in the 1955-65 birth cohorts compared to those born during the previous decade. Have neighbourhoods become less important as determinants of socioeconomic success over time or are the estimates driven by other features of the data?

First, the neighbourhood boundaries changed from 1960 to 1970. 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 correlations to increase. Second, as our estimates represent upper bounds on the neighbourhood effects, the sorting of families into local communities may have become less severe over time. If so, our estimates for the younger cohorts are simply closer to the true neighbourhood effects. The problem is time-variant sorting on the basis of unobserved family background. Assuming that the trend in sorting on the basis of observed family background is the same as on unobservables, we can check this explanation by looking at how parental education is distributed within and between neighbourhoods. If sorting has decreased from 1960 to 1970, one would expect the between-neighbourhood component (i.e. variance of the neighbourhood dummies) of the total variance in parental schooling to fall. Table 12 shows the opposite. Consequently, weaker sorting does not seem to be the explanation.

[Table 12 about here.]

Third, community misclassification error may explain the drop in neighbourhood correlation if family mobility increased during the 1950s and throughout the 1970s. We only observe location at one point in time. Since families move, 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, Page, and Solon 2001) 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 incomparable over time. Using the subset of comparable geographical units described in Section 3.1, we try to examine differential effects of staying in and moving from neighbourhood from 1960 to 1970. We limit ourselves to those aged 0–5 in 1960, simply because we want to trace the same individuals ten years later. As we discussed in section 3.1, these tract-aggregates contain more individuals and they cover larger geographical areas than the original units.

[Table 13 about here.]

The first column in Table 13 reports the unadjusted neighbour correlations for the 1955-60 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 (same individuals). As expected, aggregation implies that correlations drop, but not by much. The earnings correlations are lower for movers than for stayers, suggesting that misclassification does create a downward bias. For education, there is no significant difference between stayers and movers, although the point estimate is, surprisingly, higher for females who move than for those who stay.

Since changing neighbourhood boundaries, less sorting of families into neighbourhoods or misclassification of neighbourhoods seems to explain the drop in neighbour correlations over time, we are confident that the true impact of location during childhood has declined. 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.

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 2001). 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 (Falch and Tovmo 2001). 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 (Falch and Tovmo 2001). 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.

The school reforms implemented in the 1960s and 1970s are particularly interesting as a possible explanations for the drop in neighbourhood effect for the cohorts growing up about 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. 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 (table 190 Statistics Norway 2001).

Probably the most extensive of the reforms was the comprehensive school reform between 1960 and 1970. The aims of the reform were stated explicitly in several governmental background papers. They were to raise the level of education, to smooth the transition into higher education and to enhance equality of opportunities across socio-economic and geographical backgrounds. This was accomplished by increasing the minimum level of schooling from 7 to 9 years, unifying the education system and providing a common curriculum for all schools and by securing that there were comprehensive schools in all municipalities. It is expected that this reform reduced the effect of family background as well as neighbourhoods from 1960 to 1970, first of all since the cohorts born between 1945-55 were only partly affected by this reform and the following reforms, while the cohorts born between 1955-65 were all affected. Analyses of the participation rate to higher education for cohorts born from 1942-1970 show a strong degree of regional equalisation over time, indicating a weakening of neighbourhood background for school choice (Hægeland et al. 1999). A study focusing on the effect of the comprehensive school reform also finds that socio-economic background is weakened in terms of transition to higher education of the reform (Aakvik et al. 2001).

Access to student grants and loans was expanded in the late 1960s and early 1970s. A grant for students above 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. For practical purposes, 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.[16]

In summary, the various educational reforms did affect the postwar birth

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

cohorts differently. While the oldest cohorts typically entered upper secondary and tertiary education before these changes were fully implemented, those born after 1955 benefited from the more favourable system which basically lowered the costs of educational investment for adolescents in families located far away from relevant institutions. The more recent cohorts also experienced a system with far fewer differences in primary school organisation across local communities.

If the impact of neighbourhoods declined, one might wonder why the sibling-correlation did not drop correspondingly. But such a drop need not be a necessary corollary. When fewer families are financially constrained when making human investments, other mechanisms which tend to create homogenous 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.

5 Concluding Remarks

Family background and childhood neighbourhoods play an important role in explaining adult achievement and thus intergenerational mobility. Most studies evaluating the combined effects of family and neighbourhoods are from the US. However, it is expected that institutional differences are important determinants of the degree of equality of opportunity. In this paper we use very detailed census data from Norway which enables us to construct complete neighbourhoods and use a detailed set of family background variables. We focus especially on whether there were changes in the impact of family and neighbourhoods between 1960 and 1970. Our main results can be summarised as follows.

The sibling correlations capturing both measured and unmeasured family and neighbourhood characteristics that are shared by siblings were estimated to 0.43 for brothers and 0.47 for sisters in education (years of schooling). These figures are stable over time comparing the 1945-55 with the 1955-65 birth cohorts. The correlations are just slightly lower than those found for the United States.

Sibling correlations in permanent log earnings are around 0.16-0.21. The correlations are higher for brothers than for sisters, but the gender difference is declining over time. We find weaker effects of families on adult permanent earnings in Norway than existing US estimates, adding to the evidence suggesting that intergenerational mobility is higher in the Scandinavian welfare states than in the United States, see Björklund and Jäntti (1997) and Björklund et al. (2001).

Neighbourhood correlations in education in 1960 are 0.12 for boys and 0.109 for girls, and log earnings correlations are estimated to be 0.068 and 0.033, respectively. Comparing the 1945-55 with the 1955-65 birth cohorts, we find a declining effect of neighbourhoods as the correlations are reduced by approximately one half.

As neighbourhood correlations are upward biased because similar families cluster in communities, we tighten the bound on the variance of neighbourhood effects by using data on observed family background. Partialling out the effects of observed family background, the correlations drop considerably, for education down to 0.043 and 0.017 for the oldest boys and girls, respectively. Earnings correlations among neighbouring children are reduced to 0.054 and 0.021, for boys and girls, respectively. Even if neighbour correlations drop in both birth cohort groups, the resemblance in adult outcome is reduced over time.

We check whether the decline in neighbourhood effects can be explained by changes in neighbourhood boundaries, reduced sorting of families into communities or misclassification errors. Neither of these explanations seem plausible.

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 and those born after 1955 faced lower costs of educational investment than those born during the previous decade.

In order to give policy relevant advice, we need a better understanding of why it is that neighbourhoods seem to matter. Our aim in future research is to contribute by studying the impact of what is probably the most important neighbourhood institution: the primary school. By adding school identifiers and information about school resources, including teacher characteristics, to the data in this study, we will hope to improve our understanding of how primary schools have long-term effects on the lives of their pupils.

References

Aaberge, R., Björklund, A., Jäntti, M., Pedersen, P. J., Smith, N., and Wennemo, T. (2000). 'Unemployment shocks and income distribution: How did the Nordic countries fare during their crises?', *Scandinavian Journal of Economics*, vol. 102, pp. 77–99.

Aakvik, A., Salvanes, K. G., and Vaage, K. (2001). 'Measuring the effect of a school reform on educational attainment and earnings', mimeo, Norwegian School of Economics and Business Administration, Bergen.

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. (2001).

'Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States', forthcoming *Journal of Population Economics*.

Björklund, A. and Jäntti, M. (1997). 'Intergenerational income mobility in Sweden compared to the United States', *American Economic Review*, vol. 87, pp. 1009–1018.

Black, S. E. (1999). 'Do better schools matter? Parental valuation of elementary education', *Quarterly Journal of Economics*, vol. 114, pp. 577–599.

Borge, L.-E. and Rattsø, J. (2001). 'Spending growth with vertical fiscal imbalance: Decentralized government spending in Norway 1880–1990.', mimeo, Norwegian University of Science and Technology, Trondheim.

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, pp. 77–105.

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

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

Conley, T. G. and Topa, G. (2001). 'Socio-economic distance and spatial patterns in unemployment', forthcoming *Journal of Applied Econometrics*.

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

Ermisch, J. and Francesconi, M. (2001). 'Family matters: Impacts of family background on educational attainments', *Economica*, vol. 68, pp. 137–156.

Falch, T. and Tovmo, P. (2001). 'Norwegian local public finance in the 1930s and beyond', mimeo Norwegian University of Science and Technology, Trondheim.

Gibbons, S. and Machin, S. (2001). 'Valuing primary schools', Presented to Conference on The Economics of Education, Bergen, May 2001.

Gibbons, S. (2001a). 'Neighbourhood effects on educational achievement: Evidence from the Census and National Child Development Study', mimeo.

Gibbons, S. (2001b). 'Paying for good neighbours: Neighbourhood deprivation and the community benefits of education', mimeo.

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, pp. 603–42.

Griliches, Z. (1979). 'Sibling models and data in economics: Beginnings of a survey', *Journal of Political Economy*, vol. 87, 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, pp. 555–576.

Jargowsky, P. A. (1996). 'Take the money and run: Economic segregation in U.S. metropolitan areas', *American Sociological Review*, vol. 61, pp. 984–998.

Kahn, L. M. (1998). 'Against the wind: Bargaining recentralisation and wage inequality in Norway 1987-91', *The Economic Journal*, vol. 108, pp. 603–645.

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, pp. 607–654.

Kremer, M. (1997). 'How much does sorting increase inequality?', *Quarterly Journal of Economics*, vol. 112, pp. 115–139.

Kunz, J., Page, M. E., and Solon, G. (2001). 'Are single-year measures of neighborhood characteristics useful proxies for children's long-run neighborhood environment?', mimeo.

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.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*, Harvard University Press.

Page, M. E. and Solon, G. (2000). 'Correlations between brothers and neighboring boys in their adult earnings: The importance of being urban', Working Paper.

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, pp. 383–392.

Solon, G. (1999). 'Intergenerational mobility in the labor market', in 'Handbook of Labor Economics', (eds. O. Ashenfelter and D. Card), vol. 3a, chap. 29, Amsterdam: North-Holland, pp. 1761–1800.

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

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

Topa, G. (2001). 'Social interactions, local spillovers and unemployment', *Review of Economic Studies*, vol. 68, pp. 261–295.

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', Rapport 2/87, Statistisk Sentralbyrå.

Table 1: **Percentages of individuals linked to census data**

age-group	N	% linked
25–29	314718	91.48
30–34	294980	92.72
35–39	287952	96.19
40–44	281099	94.80
45–49	288127	90.53
50–54	232677	77.55
55–59	177185	44.50
60–64	153528	18.99
65–69	148264	10.51
70–74	130493	6.52

Note: Based on individuals from the 1995 register files.

Table 2: **Comparison of family structure**

	1920–24		1925–29		1930–34		1935–39	
	register	census	register	census	register	census	register	census
1	33.58	41.42	20.60	28.93	13.05	18.60	11.03	15.90
2	36.43	31.98	36.33	33.08	33.50	33.60	34.11	35.08
3	16.09	16.12	24.84	21.82	29.14	26.77	31.46	29.42
4	7.42	6.61	11.53	10.08	14.96	13.10	15.40	13.24
5	2.19	2.40	4.29	3.85	5.93	5.06	5.37	4.38
6	0.82	0.89	1.54	1.41	2.19	1.82	1.69	1.30
7	0.28	0.33	0.52	0.50	0.71	0.62	0.58	0.42

Note: Data from the register data on number of children compared with the size of our constructed families, for four cohorts of mothers: those born between 1920–24, 1925–29, 1930–34 and 1935–39. The percentages of mothers with the corresponding number of children is shown.

Table 3: **Comparison of samples with population**

	1945-55 cohorts			1955-65 cohorts		
	earnings	education	pop.	earnings	education	pop.
Male:						
mean age	44.26	44.49	44.54	34.44	34.31	34.43
mean of education	11.65	11.56	11.63	11.91	11.85	11.92
(standard deviation)	2.89	2.94	2.92	2.31	2.32	2.32
mean earnings 1995	269000	256000	258000	246000	231000	231000
(standard deviation)	219000	229000	350000	165000	163000	163000
full-time working	0.742	0.710	0.710	0.754	0.710	0.706
mean months unemployed	0.301	0.365	0.347	0.397	0.512	0.502
number of people	88545	120916	313629	111174	141473	328830
Female:						
mean age	44.13	44.46	44.51	34.49	34.33	34.44
mean of education	11.22	11.12	11.09	11.91	11.80	11.85
(standard deviation)	2.64	2.70	2.62	2.31	2.31	2.30
mean earnings 1995	162000	151000	150000	151000	137000	137000
(standard deviation)	90600	96100	97400	85300	91300	94500
full-time working	.475	.444	.440	.451	.410	.411
mean months unemployed	.247	.291	.286	.389	.439	.440
number of people	62905	97063	291438	93193	129657	307497
All:						
mean age	44.15	44.34	44.52	34.46	34.34	34.43
mean of education	11.50	11.40	11.37	11.94	11.86	11.89
(standard deviation)	2.78	2.82	2.79	2.31	2.31	2.31
mean earnings 1995	221000	209000	206000	201000	184000	186000
(standard deviation)	353000	326000	267000	136000	140000	142000
full-time working	.615	.585	.580	.609	.556	.563
mean months unemployed	.273	.311	.318	.386	.499	.472
number of people	262321	334597	605067	342419	455642	636327

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

Table 4: **Observations of parents**

	1945-55 cohorts		1955-65 cohorts	
	mother	father	mother	father
Years of education	7.81	8.56	8.33	9.15
standard dev. of education	1.55	2.54	1.82	2.70
Income	n.a.	n.a.	14200	35100
standard dev. of income			6375	19114
indicator of missing income			2.2%	4.8%
Separated	1.0%		2.6%	
Divorced	0.6%		2.5%	

Note: The complete set of parents that can be merged with the administrative data. Income is in current prices.

Table 5: **Possible weighting strategies**

		1	2	3	4
siblings	W_{cf}	1	S_{cf}	$\sqrt{S_{cf}(S_{cf}-1)/2}$	$S_{cf}(S_{cf}-1)/2$
	W_c	1	$\sum_{f=1}^{F_c} W_{cf}$	$\sum_{f=1}^{F_c} W_{cf}$	$\sum_{f=1}^{F_c} W_{cf}$
neighbours	$W_{cff'}$	1	$\cdot$	$\sqrt{S_{sf}S_{sf'}}$	$S_{sf}S_{cf'}$
	W_c	1	$\cdot$	$\sum_{f\neq f'} W_{cff'}$	$\sum_{f\neq f'} W_{cff'}$

Note: These schemes are taken from (Solon et al. 2000).

Table 6: **Sibling and neighbour correlations**

		1945-55 cohorts		1955-65 cohorts	
		siblings	neighbours	siblings	neighbours
education					
	male	0.4315	0.1209	0.4202	0.0613
		(0.0050)	(0.0209)	(0.0045)	(0.0048)
	female	0.4770	0.1086	0.4733	0.0618
		(0.0046)	(0.0226)	(0.0047)	(0.0053)
	all	0.4265	0.1188	0.4208	0.0607
		(0.0030)	(0.0196)	(0.0027)	(0.0047)
earnings					
	male	0.2112	0.0677	0.1749	0.0315
		(0.0058)	(0.0097)	(0.0059)	(0.0045)
	female	0.1569	0.0325	0.1557	0.0193
		(0.0067)	(0.0047)	(0.0047)	(0.0026)
	all	0.1442	0.0443	0.1221	0.0149
		(0.0030)	(0.0062)	(0.0023)	(0.0011)

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 1990–95, dropping those years before completion of education or with less than NOK (1998) 10,000 in earnings.

Table 7: **Adjusted neighbour correlations in educational attainment**

		no adj.	E	D	$E+D$	$E+D+I$
1945-55 cohorts						
	male	0.1209	0.0544	0.1196	0.0428	
		(0.0222)	(0.0137)	(0.0216)	(0.0116)	
	female	0.1086	0.0291	0.1074	0.0169	
		(0.0256)	(0.0091)	(0.0251)	(0.0074)	
	all	0.1188	0.0435	0.1181	0.0338	
		(0.0205)	(0.0086)	(0.0201)	(0.0075)	
1955-65 cohorts						
	male	0.0613	0.0256	0.0605	0.0217	0.0183
		(0.0052)	(0.0033)	(0.0052)	(0.0036)	(0.0032)
	female	0.0618	0.0203	0.0605	0.0146	0.0110
		(0.0052)	(0.0039)	(0.0060)	(0.0040)	(0.0037)
	all	0.0607	0.0219	0.0598	0.0180	0.0140
		(0.0040)	(0.0030)	(0.0050)	(0.0029)	(0.0029)

Note: The first column repeats the unadjusted correlations, the second is corrected for mother's and father's education (using a 4thd degree polynomial in parental education with first degree interactions), the third column is adjusted for the number of children in the family and dummies for seperated and divorced parents, the fourth combines the educational and demographical adjustments, and the last column includes these adjustments together with mother's and father's income, which we have for the 1970 census. The sample consists of those between 5–15 years of age in the year of the census.

Table 8: **Adjusted neighbour correlations in adult earnings**

		no adj.	E	D	$E+D$	$E+D+I$
1945-55 cohorts						
	male	0.0677	0.0566	0.0673	0.0538	
		(0.0088)	(0.0085)	(0.0098)	(0.0082)	
	female	0.0325	0.0230	0.0322	0.0205	
		(0.0049)	(0.0034)	(0.0045)	(0.0034)	
	all	0.0443	0.0333	0.0441	0.0314	
		(0.0060)	(0.0050)	(0.0060)	(0.0046)	
1955-65 cohorts						
	male	0.0315	0.0282	0.0313	0.0276	0.0256
		(0.0039)	(0.0039)	(0.0042)	(0.0037)	(0.0048)
	female	0.0193	0.0134	0.0189	0.0120	0.0104
		(0.0024)	(0.0023)	(0.0022)	(0.0022)	(0.0020)
	all	0.0149	0.0105	0.0147	0.0096	0.0076
		(0.0011)	(0.0011)	(0.0010)	(0.0010)	(0.0011)

Note: The first column repeats the unadjusted correlations, the second (E) is corrected for mother's and father's education (using a 4th degree polynomial in parental education with first degree interactions), the third column is adjusted for the number of children in the family and dummies for seperated and divorced parents (D), the fourth combines these two, and the last column also includes parental income (I) at the time of the census which is available for 1970. The sample consists of those between 5–15 years of age in the year of the census.

Table 9: **Upper bounds on the standard deviation of neighbourhood effects**

	1945–55 cohorts		1955–65 cohorts	
	male	female	male	female
adult earnings, log units	0.132	0.090	0.088	0.068
years of education	0.604	0.342	0.316	0.241

Note: Calculated using $\mathrm{sd}(\beta' Z_c) \leq \sqrt{\mathrm{cov}_{adj.}(y_{cfs}, y_{cf's'}|\tilde{X})}$.

Table 10: **Region-adjusted sibling and neighbour correlations**

		1945-55 cohorts		1955-65 cohorts	
		siblings	neighbours	siblings	neighbours
education					
	male	0.4217	0.0890	0.4167	0.0559
		(0.0050)	(0.0128)	(0.0045)	(0.0047)
	female	0.4694	0.0745	0.4685	0.0580
		(0.0047)	(0.0105)	(0.0049)	(0.0046)
	all	0.4170	0.0802	0.4165	0.0571
		(0.0031)	(0.0092)	(0.0027)	(0.0039)
earnings					
	male	0.1875	0.0320	0.1605	0.0152
		(0.0054)	(0.0039)	(0.0058)	(0.0022)
	female	0.1483	0.0108	0.1493	0.0114
		(0.0068)	(0.0024)	(0.0047)	(0.0017)
	all	0.1340	0.0178	0.1167	0.0096
		(0.0029)	(0.0027)	(0.0023)	(0.0008)

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 1990–95, dropping those years before completion of education or with less than NOK (1998) 10,000 in earnings. First-step regression includes childhood county dummies.

Table 11: **Adjusted neighbour correlations in adult earnings, region adj.**

		no adj.	E	D	$E+D$	$E+D+I$
1945–55 cohorts						
	male	0.0320	0.0264	0.0319	0.0257	
		(0.0048)	(0.0046)	(0.0048)	(0.0042)	
	female	0.0108	0.0058	0.0107	0.0052	
		(0.0020)	(0.0020)	(0.0021)	(0.0021)	
	all	0.0178	0.0124	0.0178	0.0120	
		(0.0024)	(0.0020)	(0.0025)	(0.0020)	
1955–65 cohorts						
	male	0.0152	0.0127	0.0151	0.0123	0.0110
		(0.0022)	(0.0021)	(0.0023)	(0.0020)	(0.0023)
	female	0.0114	0.0071	0.0112	0.0063	0.0054
		(0.0014)	(0.0015)	(0.0016)	(0.0013)	(0.0015)
	all	0.0096	0.0063	0.0094	0.0058	0.0046
		(0.0007)	(0.0008)	(0.0008)	(0.0007)	(0.0007)

Note: The first column repeats the unadjusted correlations, the second is corrected for mother's and father's education (using a 4th degree polynomial in parental education with first degree interactions), the third column is adjusted for the number of children in the family and dummies for seperated and divorced parents, the fourth combines the educational and demographical adjustments, and the last column includes these adjustments together with mother's and father's income, which we have for the 1970 census. The sample consists of those between 5–15 years old in the year of the census. The initial regression to produce the y_{cfs} residuals includes dummies for childhood county. Predictions on family background are on within-region variation only.

Table 12: **Degree of neighbourhood sorting**

	mother's education		father's education	
	1945-55	1955-65	1945-55	1955-65
mean	8.005	8.679	8.771	9.503
$\widehat{\sigma}_u$	0.611	0.846	0.780	1.314
$\widehat{\sigma}_\epsilon$	1.578	1.814	1.873	2.505
$\widehat{\rho} = \widehat{\sigma}_u^2/(\widehat{\sigma}_u^2 + \widehat{\sigma}_\epsilon^2)$	0.130	0.179	0.171	0.216

Note: Decomposition of the variance of parental schooling. Estimates from the fixed-effect regression $E_{ic} = \overline{E} + u_c + \epsilon_{ic}$ (neighbourhood fixed effects). Sample is restricted to parents aged 30–50 at the time of the censuses.

Table 13: **Stayers and movers**

		all (tracts)	all (aggregations)	stayers	movers
education					
	male	0.0563	0.0476	0.0497	0.0479
		(0.0044)	(0.0059)	(0.0054)	(0.0080)
	female	0.0485	0.0438	0.0418	0.0521
		(0.0035)	(0.0038)	(0.0034)	(0.0074)
	all	0.0563	0.0476	0.0497	0.0479
		(0.0047)	(0.0067)	(0.0057)	(0.0064)
earnings					
	male	0.0393	0.0368	0.0501	0.0252
		(0.0028)	(0.0041)	(0.0065)	(0.0050)
	female	0.0197	0.0196	0.0243	0.0119
		(0.0018)	(0.0019)	(0.0025)	(0.0034)
	all	0.0145	0.0125	0.0139	0.0115
		(0.0013)	(0.0011)	(0.0014)	(0.0018)

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 4969 tract aggregations with fewer than 4000 inhibitants and not containing 1960 tracts that were split among several tracts in the 1970 census. The first column summarizes 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" had moved out.