Tracking and Human Capital Inequalities: The Impact of the Finnish Comprehensive School Reform

September 12, 2014

1 Introduction

In many western countries, human capital outcomes in terms of education and health are positively associated with parental socioeconomic status (e.g. Case et al., 2002). Educational policies often aim to promote equal opportunities for all children, and one of the educational policy levers is the tracking age into differing-ability schools (Hanushek and Wößmann, 2006). In this paper we exploit a Finnish comprehensive schooling reform, which was mainly motivated by a desire to promote more equal educational opportunities, irrespective of social background (Kerr et al., 2013), to test whether delayed tracking has affected human capital disparities with respect to family socioeconomic status.

The Finnish comprehensive schooling reform in the 1970s delayed the age at which pupils were tracked into either a vocational education or a pre-academic track, from 11 years before the reform to 16 years after the reform. Late tracking means that high-performing and low-performing pupils are in the same classroom until an older age. This could benefit (i) weaker pupils who benefit from their interaction with stronger pupils, and (ii) late bloomers whose true potential reveals itself at an older age. However, the one-size-fits-all comprehensive education means that (i) both weak and strong pupils are not taught at the appropriate level – which could hamper their progress – and (ii) high-potential students are possibly held back by their weaker classmates.

Institutions seem to matter: Hanushek and Wößmann (2006) use a cross-country comparison of internationally comparable student test scores in developed countries, and report that late tracking has positive effects throughout the grade distribution, but reduces inequality by benefiting weak students more than strong students. Duflo et al. (2011) find that in developing countries, teachers have incentives to teach at the level of their top students, which implies that any positive peer effects of late tracking for weak students are offset by the negative effect of a too high level of instruction.

The Finnish schooling reform that we exploit in this paper was gradually implemented in six geographical regions between 1972 and 1977. The variation in educational policy over time and across regions allows us to cleanly identify the effect of the reform, controlling for cohort- and region-specific effects. Further, we observe an eleven percent sample of the Finnish population which includes an unusually rich set of information on childhood circumstance, educational attainment, earnings, medicine use, hospital visits, and mortality. Combining the richness of the data with a clean identification strategy, this paper is able to investigate whether educational policies can influence socioeconomic disparities in human capital (that is, educational attainment and health).

Using a similar research design, previous studies have found that the Finnish compulsory school-ing reform lead to an increase in the gender differences in educational attainment, to a decrease of the gender wage gap in adult income by four percentage points Pekkarinen (2008), to a decrease in the intergenerational income elasticity Pekkarinen et al. (2009) and to an increase in the military test scores of boys from parents with low levels of education Kerr et al. (2013).

In Sweden, Meghir et al. (2013) exploited a schooling reform in the 1950s and early 1960s that mainly increased the compulsory years of schooling by one or two years depending on municipality. They find a positive effect on military test scores, with the strongest gains for boys from low-income families and of low ability. Moreover, they find that the reform increased the completed years of schooling of boys with low-educated fathers by almost three and a half months. While the authors do find a decline in mortality for low-ability individuals, they conclude that the reform only had a minor effect negative effect on the number of sick days of men with low-educated fathers, and they find no effect on hospitalization.

Jones et al. (2014) relate childhood circumstances to later-life health outcomes using the UK National Child Development Study (NCDS), and explicitly frame their empirical analysis into the inequality of opportunity framework (e.g. Roemer, 1998). The authors then exploit that members of their 1958 cohort were exposed to different educational regimens – a comprehensive schooling reform in the 1960s affected only some part of the cohort – to study the effect of different educational policies on inequality of opportunity in health. The results are mixed: the worst-off groups in society are better off in terms of some health outcomes under the new comprehensive schooling system, while they are worse off in terms of other health outcomes.

Our contribution to the literature is the investigation of the effects of the tracking age on human capital disparities with respect to household income during childhood. Compared to the series of papers of Pekkarinen and co-authors on the Finnish reform, we are the first to focus on disparities in both health and education with respect to household income during childhood. Our contribution over Meghir et al. (2013) and Jones et al. (2014) is that we are able to cleanly identify the effect of a reform that changed the tracking age. In contrast, Jones et al. (2014) only compare individuals in different schooling systems, but do not estimate the effect of the schooling reform, and the Swedish reform does not allow for a separate estimation the effect of increased years of schooling and tracking. Moreover, the Finnish reform was gradually implemented across six regions, resulting in six clusters for computing the standard errors of our estimates. This gives us a competitive edge in terms of statistical precision compared to the Swedish case, where for some results the appropriate number of clusters is only two, leading to a huge drop in precision.[1]

We find evidence that delaying the educational tracking age in Finland in the 1970s had a per-sistent impact on human capital throughout the life-cycle. First, the reform has increased longevity for women, which demonstrates that late-childhood interventions can have a lasting impact through-out adulthood. Second, the reform generally reduced disparities with respect to household income during childhood: in terms of educational attainment for women, and in terms of accident risk hypertension for men.[2] Yet, we find that the tide of the Finnish compulsory schooling reform did not lift all boats. While the reform reduced socioeconomic disparities, the gains for children from poor families seem to have come at the cost of children from affluent families.

This paper is organized as follows. In the next section, we outline the theoretical mechanisms by

[1]The results of Meghir et al. (2013) for test scores and mortality are based on variation between the 1948 and 1953 birth cohorts and between reformed and unreformed municipalities. Their results for educational attainment, hospitalization and sick leave are based on variation between municipalities and all birth cohorts between 1946 and 1957. A potential limitation of their approach is that standard errors are clustered at the municipality level, which could lead to over-rejection of null hypotheses due to underestimated standard errors and wrong threshold values of the t-distribution. Standard errors should be clustered at the level of assignment of the treatment, which by definition cannot exceed the number of years in which treatment assignment changed. In the case of military test scores and the significant results for mortality for low-ability men, two clusters – not 1,055 as the authors claim – are appropriate, while the appropriate number of clusters for the remaining analyses is twelve. In the case of Finland, where the reform was gradually implemented between 1972 and 1977, we cluster at the level of six regional reform regions, instead of clustering at the level of the municipality of residence which would lead to over-rejection.

[2]There is however suggestive evidence at a 10% level of significance that the reform increased disparities with respect to household income for antidepressants for girls.

which delayed tracking could affect human capital decisions and mortality. In section 3 we discuss the comprehensive schooling reform, and in section 4 the data. Section 5 discusses the empirical methodology, section 6 the results, after which we conclude in section 7.

2 Socioeconomic disparities and the optimal tracking age

Cunha et al. (2006) outline a model of human capital formation where the effect of an early life intervention is multiplied by interactions with later life investments: the intervention sets off a chain of positive events throughout the life course. Despite empirical evidence in favor of this hypothesis for the affected individuals (e.g. Campbell et al., 2014), microeconomic evidence does not necessarily lead to the right macroeconomic implications in the presence of spillover effects. For example, the role of negative spillover effects has been documented for the labor market, where job search assistance for some job seekers leads to displacement of other workers Cr´epon et al. (2012); Gautier et al. (2012). In the case of educational policy, since school grades are often rank-based, interventions that particularly benefit students with a disadvantaged background may “push downwards” the grades of their classmates, thereby adversely affecting their employment prospects. This type of reasoning does not solely apply to grades. In the remainder of this section, we outline two further mechanisms through which the effect of delayed tracking, as in the Finnish schooling reform, can have positive effects on the development and performance of some students, and negative effects for others: through peer effects and the “distance” between the level of the student and the level of instruction.

Student performance at age t is determined by (i) ability, assumed to be largely constant over time, (ii) childhood circumstance, which is assumed to have a strong effect at younger ages that gets weaker as the child gets older, (iii) peer effects, which are increasing in the performance of the student’s classmates, and (iv) the distance between past student performance and the level of instruction, which is determined by the average or median student performance. Up to the tracking age, high- and low-performing students are in the same classroom in comprehensive school with a one-size-fits-all curriculum. They are tracked into either an academic or a vocational track, based on their performance.

Given these four determinants of student performance, a student at a younger age may not have revealed his true adult potential. Instead, the student’s performance is codetermined by child-hood circumstance which by assumption becomes less important as the student gets older. Hence, tracking at a young age based on student performance depends greatly on childhood circumstance, and may lead to mistracking: as the adult potential of a student manifests itself in late childhood, it may become apparent that the student is in a track that is either too demanding or too easy. For a student who is undertracked, his/her peers may drag down the performance and the level of instruction may be too low. By the end of their schooling career, they have academically suffered from interaction with weaker peers and a low level of instruction. A student who is overtracked benefits from the interaction with his/her higher ability peers, but may suffer from the advanced level of instruction and may fail the exams. Switching tracks midway involves switching costs – the costs of adjusting to a new learning environment.

The upside of early tracking, as in the old Finnish schooling system, is that students – at least those who were not mis-tracked – are taught at an appropriate level which matches their own performance. Students in the high-level track are not held back by their low-achieving peers. However, students in the low-level track are not able to benefit from high-achieving classmates.

Late tracking, as in the new Finnish schooling system, means that students with high and low performance remain in the same classroom until an older age when their performance more accurately reflects ability, such that mis-tracking is less likely. While low-performing students benefit from interacting with their high-performing peers, high-performing students may suffer from interacting with weaker students. Moreover, both high-achieving and low-achieving classmates are taught at the same level of instruction, which potentially hampers their development.

The preceding discussion shows that we expect heterogeneous effects of the Finnish compulsory schooling reform, depending on ability and childhood circumstance. The empirical test here is to investigate which of the theoretical mechanisms – the likelihood of mis-tracking, peer effects, and the distance between the level of instruction and the level of the student – dominate by determining the sign of the treatment effect. While we do not observe childhood ability, we do observe parental socioeconomic status, and so we will allow the treatment effect of the reform to vary with parental socioeconomic status of the child.

3 The Finnish comprehensive school reform

The Finnish comprehensive school reform was gradually implemented between 1972 and 1977. Finnish children start primary school at the age of 7 (see figure 1) but before the reform, children were separated by academic ability into one of two different tracks at age

11. Tracking was based on an entrance examination, school grades and teacher assessment. The high track prepared for upper secondary education and university while the low track prepared for civic school and vocational school. After the comprehensive school reform, tracking was postponed until the age of 16. It meant that children of different academic ability were held together in the same classes throughout the 9 years of comprehensive education.

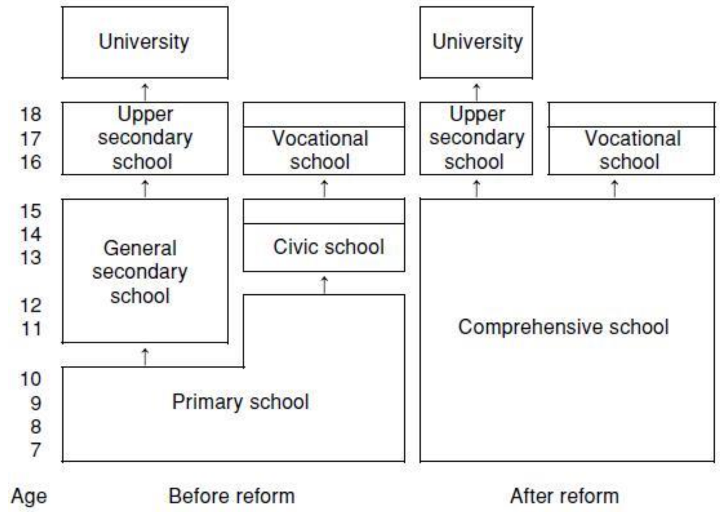


Figure 1: Finnish school system before and after the reform. Source: Pekkarinen 2009.

The reform was gradually implemented from Northern to Southern Finland in six implementation regions in the six years between 1972 and 1977. . It was somewhat controversial because it raised concerns that education in the new comprehensive school would be of lower quality than in the old general secondary school. Opposition was most fierce in the densely populated Helsinki region, and figure 2 shows that Helsinki was the last region to implement the reform. In the year of the reform, pupils who were already in grades 6 through 9 continued in the old system, while students starting grades five and below were transferred to the new system. To illustrate this, we look at the first implementation region, where the reform was implemented in 1972. The 1961 birth cohort that started the 5th grade in the year of the reform went through the old system for 4 years and then through the new system for 5 years. Pupils from this region who were born in 1965 were the first to only attend comprehensive school. This implementation scheme is illustrated in table 1.

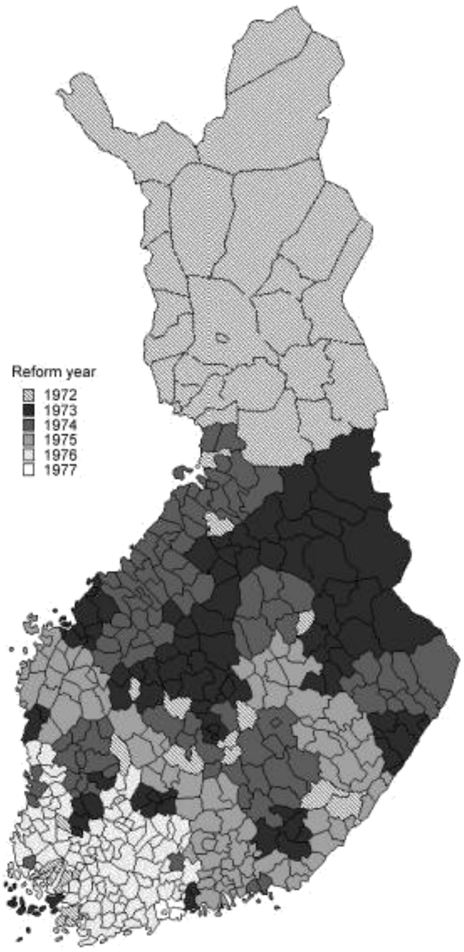


Figure 2: The year of the implementation of the reform differed between municipalities. Source: Pekkarinen 2009.

		Region 1972	Region 1973	Region 1974	Region 1975	Region 1976	Region 1977
Birth cohort	1960	C	C	C	C	C	C
	1961	T	C	C	C	C	C
	1962	T	T	C	C	C	C
	1963	T	T	T	C	C	C
	1964	T	T	T	T	C	C
	1965	T	T	T	T	T	C

Table 1: Birth cohorts by region under the old regime (controls, "C") or in the new system (treatment cells, "T"). Region 1972 refers to the region that underwent the reform in 1972, region 1973 refers to the region that underwent the reform in 1973, and so on.

4 Data on socioeconomic status, education, and health in Finland

We use an 11 percent sample of Finnish residents who were born between 1959 and 1966 and who were residing in Finland between 1987 and 2007. The total sample size for our main analysis is 56,502 individuals. We will analyse males and females separately to allow for a differential response to the schooling reform by gender. In this section, we will discuss our variables relating to family socioeconomic status, the individuals educational attainment, and individuals health outcomes, respectively.

Family socioeconomic status

Through a family identification number, we can link children to their parents and observe parental characteristics such as income and educational attainment. Information on income is obtained through the tax register, and our main measure of family socioeconomic status is the average of household incomes observed in 1970 and 1975 (i.e. during the childhood of the individual). Table 2 shows sample characteristics for males and females in the the 1960 to 1965 birth cohorts. Household income was equal for boys and girls, and the high inflation level led to a dramatic increase in nominal household income. Figure 3 shows the distribution of household income during childhood for the 1960 to 1965 birth cohorts, corrected for inflation. The hump-shaped right tail is due to top-censoring of income at 17,705 in 1970 and at 17,800 in 1975 – both in 1975 prices.

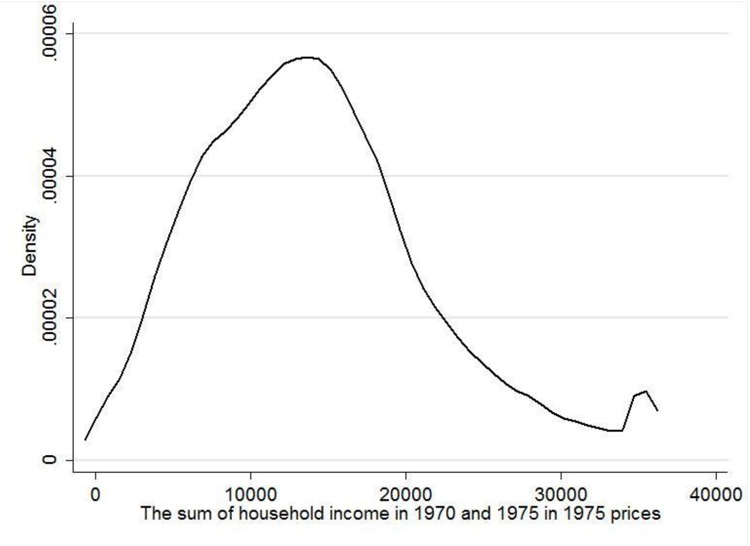


Figure 3: Probability density function of household income in 1970 and 1975 in 1975 prices.

Individuals educational attainment Retrospective information on educational attainment predating 1970, the first census year, is available from the 1980 census. Education is measured as the highest degree that was obtained, while precise information on years of schooling is not available. Table 2 shows that close to sixteen percent of men and eighteen percent of women obtained at least a college degree.

Health outcomes

In order to be able to answer our main research question, we focused on selected health outcomes that show sufficient health variability and a socioeconomic gradient by middle age, as the cohorts affected by the schooling reform were only in their mid-forties by 2007, our last observation year. Health outcome data were obtained from linking the sample to three other datasources: the mortality registry, the hospitalization registry, and a dataset with drug prescription information by diagnosis for outpatient treatment. The latter information is fairly unique as in most countries, data on chronic diseases can only be derived from the admission diagnosis of hospitalized patients, which misses out on all non-hospitalized patients. All medically prescribed drug purchases in Finland are eligible for reimbursement, and therefore there are no financial barriers for seeking pharmaceutical treatment of chronic disease (Nihtila et al., 2008).

We use four indicators for adult health, taken from the three linked data sources. (1) Survival until 2012, obtained from the cause-of-death registration, (2) the number of hospital admissions for an accident between 1987 and 2007 from the hospitalization registry, (3) total number of purchases of hypertension medication between 1995 and 2007, and (4) total number of purchases of antidepressants between 1995 and 2007. Survival, accidents, hypertension and depression are all health indicators that display a socioeconomic gradient in Finland in middle age adults. We chose a cause of significant morbidity (accidents) and two chronic conditions (hypertension and depression) that are sufficiently prevalent among young adults in order to have sufficient power to detect relevant health differences among this relatively healthy population group.

Table 2 shows that men were much more likely than women to be hospitalized for an accident, and less likely to purchase hypertension medication – 25 percent of men and 30 percent of women – and antidepressants – 20 and 27 percent, respectively – between 1995 and 2007.

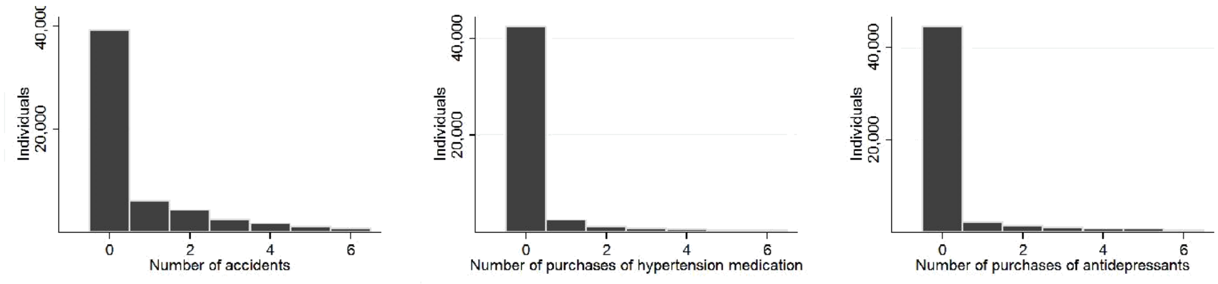
.

Figure 4 shows that the distributions of the medical variables are strongly skewed to the right, due to the large mass at zero. 26 percent have at least once been hospitalized for an accident, and the maximum number of accidents is in the data is 33. 99 percent of individuals in the sample has made less than 80 purchases of hypertension medication an 99 percent has made less than 64 purchases of antidepressant medication.

Table 2: Descriptive statistics, by gender

	Male	Female	Total
Household income in 1970	3583.2	3572.7	3578.3
	(2200.7)	(2200.3)	(2200.5)
Household income in 1975	7619.4	7646.4	7632.0
	(4227.5)	(4228.8)	(4228.1)
College degree	0.155	0.180	0.167
	(0.362)	(0.385)	(0.373)
Earnings in 2000	24595.2	18578.7	21787.7
	(13926.9)	(10623.8)	(12850.1)
Total accidents	1.576	0.640	1.139
	(3.645)	(1.787)	(2.965)
Hypertension medication	0.254	0.300	0.275
	(0.435)	(0.458)	(0.447)
Total number of purchases of hypertension medication	5.799	5.744	5.773
	(17.92)	(16.68)	(17.35)
Antidepressants	0.204	0.274	0.237
	(0.403)	(0.446)	(0.425)
Total number of purchases of antidepressants	0.769	1.140	0.942
	(2.116)	(2.616)	(2.370)
Deceased	0.0244	0.0105	0.0179
	(0.154)	(0.102)	(0.133)

Sample means, standard deviations in parentheses. Mortality for the period between 1987 and 2007, medication and accidents for the period between 1995 and 2007. Source: Statistics Finland



(a) Accidents (b) Purchases of hypertension medication (c) Purchases of antidepressants

Figure 4: Truncated distributions of the medical variables.

5 Estimation methods

We employ a variation on the differences-in-differences estimator to estimate the effect of the reform on the intergenerational income effect on health. In particular, we test whether0 the reform has changed the influence of household income during childhood on individuals' human capital outcomes later in life. Our identification strategy relies on the assumption that differences between treatment regions remain constant across the 1960 to 1965 birth cohorts: cohort effects are assumed to be the same for each of the six regions in absence of the reform.

We do this by regressing our human capital outcomes – educational attainment and health indicators – on the natural logarithm of household income during childhood h – the inflation-adjusted sum of 1970 and 1975 household income –, a treatment dummy , and the interaction between the two. We control for five birth cohort dummies relating to the 1960 to 1965 birth cohorts in vector c, five reform region dummies in vector r, and the interaction between those two sets of dummies and log household income in 1975 to account for differential effects of household income across birth cohorts and regions. The regression equation then becomes:

$$
\begin{aligned}
{}_{,,} &= {}_0 + {}_1 h_{,,} + {}_{2\,,,} + {}_3 h_{,,,,,} \\
&\quad + {}_4' \mathrm{c}_c + {}_5' \mathrm{r}_r + {}_6' \mathrm{c}_c h_{i,c,r} + {}_7' \mathrm{r}_r h_{i,c,r} + {}_{i,c,r} \qquad (1)
\end{aligned}
$$

where $_{,,}$ is the outcome variable for individual in cohort in region . $_1$ reflects the gradient between household income and the relevant outcome for the baseline cohort 1960 in the first region that experienced the reform, $_2$ reflects the treatment effect of the reform, and $_3$

reflects how the gradient has changed due to the reform. We exclude the 1966 birth cohort as all individuals in this cohort are treated and we do not have observations for an untreated region to construct the counterfactual based on the common trend across regions.

We cluster standard errors at the regional level, which is the level of variation of the treatment variable. This leads to higher threshold values for the t-statistic for each given significance level (Angrist and Pischke, 2009).

6 Results

6.1 Educational attainment

Table 3 presents the results for our first outcome measure: the probability of completing at least a college degree. The first row of the table shows the association between the natural logarithm of household income during childhood and educational attainment for the 1959 and 1960 birth cohort. On average, a one hundred percent increase in household income during childhood is associated with an eleven percent predicted increase of the probability of completing at least a college degree for both men and women.

The second and third rows of table 3 show the estimated average marginal effects of the reform based on a logistic regression version of equation 1, since the outcome we employ here is binary. All results are similar when we run a linear instead of a logistic regression model. The evidence shows that while there is no effect of the reform on higher education for individuals at the median of household income, the reform did reduce the gradient between household income and educational attainment by twenty percent. The fact that the reform reduced disparities in educational attainment by household income, suggests a possible pathway for the effect of the reform on the reduction of the intergenerational income elasticity that was reported by Pekkarinen et al. (2009).

The results for educational attainment seem to be driven mainly by the effect on women. The reform increased the probability of finishing college for girls from median income families, and reduced disparities with respect to household income for girls. However, under the maintained assumption that the effect of the treatment was linear in household income during childhood, the model predicts that while the effect for girls from the 81 percent poorest families is positive, the estimated effect for girls from the 19 percent richest families in 1970 and 1975 – those with incomes above 20,000 Euros in 1975 prices – is negative. The equalization of opportunities seems to have come at the expense of girls from the richest families: while the reform benefited most girls, it appears to have harmed girls from the most affluent backgrounds.

Table 3: Results for higher education

	Total	Men	Women
Log Parental income	0.124***	0.124***	0.125***
	(0.004)	(0.005)	(0.006)
Reform × parental income	-0.020**	-0.015	-0.028**
	(0.009)	(0.016)	(0.014)
Reform	0.001	-0.007	0.011**
	(0.005)	(0.008)	(0.005)
Observations	56502	30268	26234

The first panel refers to the association between the natural logarithm of the sum of household income (the sum of 1970 and 1975 in 1975 prices) during childhood and a binary outcome variable which indicates whether an individual has obtained at least a college degree for the 1960 and 1959 birth cohort, expressed as average marginal effects from a logit model. The coefficient of an indicator variable for the 1959 birth cohort was omitted from the table. The second panel refers to the estimated effect of the reform on a binary outcome variable which indicates whether an individual has obtained at least a college degree across the distribution of household income (the natural logarithm of the sum of 1970 and 1975 in 1975 prices) during childhood. Log earnings were normalized by subtracting the median value, such that the coefficient of the reform dummy can be interpreted as the predicted effect of the reform for children from households at the median of the income distribution. The sample is limited to the 1960 to 1965 birth cohorts. The coefficients of the cohort and region dummies and their interactions with household income are omitted from this table. Standard errors are clustered at the region level.* indicates significance at the 10 percent level, ** at the 5 percent level, and ***

at the 1 percent level.

6.2 Health outcomes

Apart from educational attainment, another critical component of human capital is health.

Mortality

Table 4 presents evidence that the effect of the comprehensive schooling reform extends over the full life course and impacts mortality. We estimate a Cox proportional hazard model for survival duration between 1987 and 2008 (i.e. between ages 27 and 48 for the 1960 birth cohort), taking into account left-truncation and right-censoring. The first row of table 4 illustrates that, on average, individuals from richer households are less likely to die. This result is mainly driven by males: men who grew up in affluent families have lower mortality rates than men who grew up in low-income families. We do not find evidence of a relationship between parental income and mortality for girls: the estimated hazard ratio for women is not significant and close to one.

Strikingly, the third row of table 4 suggests that the reform led to a drastic reduction in early-mortality for women. In contrast, for men the hazard ratio is positive, yet statistically significant at

Table 4: Duration analysis for mortality on regular sample

	Total	Men	Women
Parental income	0.811***	0.757***	0.971
	(0.0378)	(0.0392)	(0.0948)
Reform × parental income	1.032	1.148	0.715
	(0.1800)	(0.1932)	(0.1910)
Reform	1.064	1.246*	0.583***
	(0.1112)	(0.1424)	(0.1141)

Hazard ratios for mortality from a Cox proportional hazard model with left truncation in 1987 and right censoring in 2008. The first panel refers to the association between the sum of household income (the sum of 1970 and 1975 in 1975 prices) during childhood and mortality for the 1959 and 1960 birth cohorts in the regular sample. The coefficient of the 1959 birth cohort dummy was omitted from this table. The second panel refers to the estimated effect of the reform on mortality across the distribution of household income (the sum of 1970 and 1975 in 1975 prices) during childhood. The sample is limited to the 1960 to 1965 birth cohorts in the regular sample. The coefficients of the cohort and region dummies and their interactions with household income are omitted from this table. Standard errors are clustered at the region level.* indicates significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.

10% only. The second row shows that we do not find evidence that the reform affected individuals differently depending on their family background.

We now turn to the three health outcomes in our data that are associated with parental background for the 1960 birth cohorts. The zero-inflated negative binomial model seems to be appropriate.[3] This implies that for accidents, the results consist of two parts. The first part is a logit model for predicting excess zeros, and the second part is a negative binomial model for the counts greater than zero. All covariates in model 1 are included in both the negative binomial part of the model and the logistic part for predicting excess zeros.

The first row of panel A of table 5 show a strong pre-reform association between household income during childhood and accidents that required hospitalization for men but not for women. For men, a one hundred percent increase in household income is associated with a 14 percent

[3]An overdispersion test rejected the assumption of no overdispersion (p=.000) which favors the negative binomial model over the Poisson model, the Vuong test rejected the assumption of no excess zeroes (p=.000) which favors the zero-inflated negative binomial model over the regular negative binomial model, and a likelihood ratio test (p=.000) favors the zero-inflated negative binomial model over the zero-inflated poisson model due to overdispersion in the former model. A hurdle model seems inappropriate here since the processes of whether any accident-related

hospitalization occurred, and subsequently how many. are not independent of each other.

reduction in the total number of accidents between 1995 and 2007. We do not find evidence that household income is associated with the likelihood of having zero accidents (second row, panel A).

The third and fourth rows of panel A show that the reform led to a higher number of accidents for men, and that the effect was stronger for men from affluent families. However, the fifth and sixth rows indicate that the reform simultaneously increased the probability of zero accidents for men, especially those who grew up in affluent families. While these latter results are somewhat puzzling, generally the results indicate that the reform increased the number of accidents for men, in particular those from richer families, thereby reducing inequality in accidents with respect to household income.

For hypertension medication and antidepressants, we favor the negative binomial logit hurdle model over the zero-inflated negative binomial model. This is because we view the choice to ever buy medication as a separate process from the choice of how many times to purchase medication.[4] Panels B and C of table 5 refer to average marginal effects from the Logit part of the negative binomial logit hurdle model. None of the negative binomial coefficients of this model were significant: we do not find evidence of an effect on the amount of medication purchases. For clarity of exposition the coefficients of the negative binomial part of the hurdle model were omitted from the table. The results therefore should be interpreted as being on the extensive margin, i.e. the choice of buying any hypertension/antidepressant drugs versus not buying at all.

The first row of panel B of table 5 shows that parental income is negatively correlated to the likelihood of purchasing antidepressant medication for the 1959 and 1960 birth cohorts. The second and third lines show that the compulsory schooling reform reduced these disparities for boys, while our estimates for girls lack precision. We do not find evidence for an aggregate effect.

The first row of panel C of table 5 shows a negative association between parental income and antidepressant use for women. The second row of panel C shows that there is not sufficient evidence at the five percent significance level to conclude that the reform affected the socioeconomic disparities in antidepressant use, although there is suggestive evidence at a ten percent significance level that the reform increased the gradient with respect to antidepressants for girls. The third line

[4]Specification tests reject the poisson, negative binomial, and zero-inflated poisson model. Both the negative binomial hurdle model and the zero-inflated negative binomial model pass the test, yet we prefer the former for the reason given in the text.

Table 5: Results for health outcomes

	Total	Men	Women
Panel A. Total accidents			
Negative binomial coefficients			
Parental income	-0.078**	-0.141***	0.072
	(0.040)	(0.050)	(0.064)
Logit coefficients for predicting excess zeros			
Parental income	0.032	0.024	0.098
	(0.062)	(0.091)	(0.091)
Negative binomial coefficients			
Reform × parental income	0.615***	0.748***	0.141
	(0.132)	(0.199)	(0.207)
Reform	0.476***	0.483**	0.258
	(0.150)	(0.205)	(0.223)
Logit coefficients for predicting excess zeros			
Reform × parental income	0.571***	2.013***	0.050
	(0.165)	(0.568)	(0.097)
Reform	0.384**	0.752*	0.128
	(0.180)	(0.398)	(0.145)
Panel B. Hypertension medication			
Parental income	-0.027***	-0.023***	-0.031***
	(0.004)	(0.005)	(0.005)
Reform × parental income	0.013	0.038**	-0.075
	(0.013)	(0.016)	(0.110)
Reform	0.010	0.012	0.033
	(0.007)	(0.009)	(0.041)
Observations	56502	30268	26234
Panel C. Antidepressants			
Parental income	-0.015***	0.001	-0.024***
	(0.003)	(0.006)	(0.007)
Reform × parental income	0.004	0.032	-0.027*
	(0.011)	(0.022)	(0.015)
Reform	0.022**	0.032**	0.011
	(0.009)	(0.014)	(0.008)
Observations	56502	30268	26234

The first two rows of panel A refer to a zero-inflated negative binomial regression of the total number of accidents between 1995 and 2007 on the natural logarithm of the sum of household income (the sum of 1970 and 1975 incomes in 1975 prices and in the table referred to as parental income) during childhood for the 1960 and 1959 birth cohort. The coefficient of a 1959 birth cohort dummy was omitted for clarity of exposition. The final four rows of panel A refer to our treatment estimates that were obtained from a separate zero-inflated negative binomial model, the coefficients of cohort and region dummies and their interactions with parental income are not reported. The median value of parental income in the sample was subtracted for all individuals, the treatment coefficient therefore refers to the predicted effect for children from median-income parents. Panel B refers to average marginal effects from logistic regression models with having bought hypertension medication between 1995 and 2007 as the outcome variable. As in panel A, the first row of panel B shows the association for the 1959 and 1960 birth cohort, and the final two rows of Panel B refer to our estimated treatment effects from the full model. Panel C presents results from the same models as panel B, but now for a binary outcome variable that indicates whether the individual purchased antidepressants between 1995 and 2007. Standard errors are clustered at the region level.* indicates significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.

shows that the reform increased the likelihood of antidepressant use by men from median-income households, while there is no evidence that the effect was heterogeneous across the household income distribution.

7 Conclusion

Exploiting regional variation in the moment of tracking in Finland, we find that delaying the tracking age generally reduces the gradient between socioeconomic circumstance during childhood and human capital outcomes later in life. This conclusion holds for the likelihood of obtaining a higher education for girls, and for men, the reform lowered the socioeconomic disparities with respect to accidents that require hospitalization and hypertension medication use. Only, for antidepressant use among women there is some suggestive evidence that the reform may have increased the gradient. Finally, our results suggest that the reform had a positive effect on longevity for women.

This paper is the first to comprehensively study the effect of tracking on health outcomes. Health may have been affected by the reform directly, through a different set of peers and health knowledge obtained in school, but also indirectly through different career paths and income levels as a result of the reform. Apart from mortality, three of our health measures are related to health care use. Health care seeking behavior, such as medicine use, may simultaneously reflect both health and health knowledge. Our results for health care use therefore cannot separate possible effects on both health and health knowledge.

Our modified differences-in-differences framework adds to the literature by accounting for unobserved heterogeneity that remained constant across birth cohorts but varied over the six reform regions, and for unobserved heterogeneity that affected birth cohorts differently but that was constant across regions. We find evidence of effects of the comprehensive schooling reform throughout the life course on educational attainment, outpatient treatment for chronic disease, and mortality. This suggests that late tracking reduces human capital disparities with respect to parental background, but the human capital gains for children from poorer families may have partly come at the cost of children from affluent families.

References

Angrist, J. D. and Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). Early childhood investments substantially boost adult health. Science, 343(6178):1478–1485.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. The American Economic Review, 92(5):pp. 1308–1334.

Cr´epon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2012). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. Technical report, National Bureau of Economic Research.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. Handbook of the Economics of Education, 1:697–812.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. The American Economic Review, pages 1739–1774.

Gautier, P., Muller, P., Van der Klaauw, B., Rosholm, M., and Svarer, M. (2012). Estimating equilibrium effects of job search assistance. Technical report, Tinbergen Institute Discussion Paper.

Hanushek, E. A. and W¨oßmann, L. (2006). Does educational tracking affect performance and inequality? differences- in-differences evidence across countries*. The Economic Journal, 116(510):C63–C76.

Jones, A., Roemer, J., and Rosa Dias, P. (2014). Equalising opportunities in health through educational policy. Social Choice and Welfare, pages 1–25.

Kerr, S. P., Pekkarinen, T., and Uusitalo, R. (2013). School tracking and development of cognitive skills. Journal of Labor Economics, 31(3):577–602.

Meghir, C., Palme, M., and Simeonova, E. (2013). Education, cognition and health: Evidence from a social experiment. Technical report, National Bureau of Economic Research.

Nihtila, E. K., Martikainen, P. T., Koskinen, S. V., Reunanen, A. R., Noro, A. M., and Hakkinen, U. T. (2008). Chronic conditions and the risk of long-term institutionalization among older people. The European Journal of Public Health, 18(1):77–84.

Pekkarinen, T. (2008). Gender differences in educational attainment: Evidence on the role of tracking from a finnish quasi-experiment. The Scandinavian Journal of Economics, 110(4):807– 825.

Pekkarinen, T., Uusitalo, R., and Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the finnish comprehensive school reform. Journal of Public Economics, 93(7):965–973.

Roemer, J. E. (1998). Equality of opportunity. Harvard University Press.