The Impact of Later Tracking on Mortality by Parental Income in Finland

Bastian Ravesteijn[1]
Hans van Kippersluis[2]
Mauricio Avendano Pabon[3]
Pekka Martikainen[4]
Hannu Vessari[5]
Eddy van Doorslaer[6]

[1] Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands; Department of Health Care Policy, Harvard Medical School, USA; LIRAES (EA4470), Université Paris Descartes - Sorbonne Paris Cité, France
[2] Erasmus School of Economics, Erasmus University Rotterdam and Tinbergen Institute, The Netherlands
[3] Department of Social Science, Health and Medicine at King's College London, United Kingdom; Department of Social and Behavioural Sciences, Harvard School of Public Health, USA
[4] Department of Sociology, University of Helsinki, Finland
[5] Novo Nordisk, Aalto, Finland
[6] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, the Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Can education reduce health inequalities? The impact of later tracking on mortality by parental income in Finland

Bastian Ravesteijn[*,a,b,c], Hans van Kippersluis[a,d], Mauricio Avendano Pabon[e,f], Pekka Martikainen[g], Hannu Vessari[h], and Eddy van Doorslaer[a,d,i]

[a]Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

[b]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

[c]LIRAES (EA4470), Université Paris Descartes - Sorbonne Paris Cité, Paris, France

[d]Tinbergen Institute, Rotterdam, the Netherlands

[e]Department of Social Science, Health and Medicine at King's College London, London, United Kingdom

[f]Department of Social and Behavioural Sciences, Harvard School of Public Health, Boston, MA, USA

[g]Department of Sociology, University of Helsinki, Helsinki, Finland

[h]Novo Nordisk, Aalto, Finland

[i] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, the Netherlands

February 24, 2017

---

[*]Correspondence to: ravesteijn@hcp.med.harvard.edu

Abstract

We investigate whether later educational tracking reduced the intergenerational persistence of socioeconomic disparities in mortality in Finland, where the tracking age was raised from 11 to 16 in the 1970s. We use a difference-in-differences approach that exploits the gradual rollout of the reform. We find that late tracking did reduce disparities in mortality around the age of 50 by parental income for men. However, the longevity gains of men from low-income families seem to have come at the cost of increased mortality among men who grew up in high-income families. This raises questions about the welfare implications of the reform.

1 Introduction

Health has been found to be positively correlated with parental socioeconomic status, with children from low-income families reporting worse health and experiencing lower life expectancy than their high-income counterparts (e.g. Case et al. 2002, Palme & Sandgren 2008). Educational policies could promote equal opportunities for children, and one of the educational policy levers is the choice of the age at which pupils are tracked into differing-ability classrooms (Hanushek et al. 2006). Early tracking implies separating pupils at a young age into different classes according to the students' overall academic achievement, while late tracking keeps students with different levels of ability together in the same class until a later age. Late tracking could have both positive and negative effects on student outcomes, with potentially heterogeneous effects depending on parental socioeconomic status (see e.g. Betts et al. 2011, who describes the political controversy around late tracking). While later tracking might yield positive

outcomes for children from low-income families, it may come at the cost of human capital accumulation for children from high-income families, who might benefit less from a classroom of mixed academic ability. Recent empirical evidence suggests that late tracking increases human capital accumulation among children from low-income families, but there is no prior evidence on whether these effects persist over the life course and influence the ultimate health outcome: mortality.

Our main contribution is to test whether raising the tracking age from 11 to 16 in Finland has reduced later-life mortality differences between children from low- and high-income families. While previous studies have examined the long-term effect of compulsory schooling laws on health (Van Kippersluis et al. 2011, Clark & Royer 2013), no studies have explored the health impact of the tracking age, an important feature through which education systems influence social mobility. Unique is the fact that the Finnish reform changed the tracking age but had no impact on the years of schooling, enabling us to isolate the impact of tracking alone. In addition, our study breaks new ground by examining how the reform differentially affected the health of children from different socioeconomic backgrounds. In particular, we investigate whether the potential longevity gains of tracking for those who grew up in poor families came at the expense of reduced longevity for individuals who grew up in rich families.

The empirical literature suggests that educational tracking matters for educational outcomes and earnings, but the effects on health remain relatively unexplored (Grossman 2015). Previous research has found that countries with late tracking show less inequality in academic achievement outcomes (Hanushek et al. 2006, Ammermueller 2013, Woessmann 2016). The best evidence to date comes from several Northern European countries which raised the tracking age in the second half of the twentieth century. Meghir et al. (2013) evaluate a Swedish schooling reform in the 1950s and early 1960s

3

that (i) raised the compulsory years of schooling by one or two years depending on the municipality of residence, and (ii) raised the tracking age. They find a positive total effect of the compulsory schooling and tracking reform on military test scores and completed years of education, with the strongest gains for boys from low-income families and boys of low ability.

Pekkarinen and co-authors have investigated the effect of the educational reform of 1972-77 in Finland on education and labor market outcomes. The Finnish reform raised the tracking age from 11 to 16, but in contrast to the Swedish reform no impact on the average years of schooling has been demonstrated. They find that the reform led to an increase in the military test scores of sons from parents with low levels of education (Kerr et al. 2013), while leading to a four-percentage point fall in the gender wage gap by four percentage points (Pekkarinen 2008), and to a decrease in the intergenerational income elasticity for sons but not for daughters (Pekkarinen et al. 2009). Whether these effects on education and income translated into longer and healthier lives, however, has not been firmly established (Mazumder 2012, Lindahl 2005, Cesarini et al. 2016).

Only two studies have explored the effect of tracking on health. Meghir et al. (2013) found that the Swedish reform led to a reduction in health inequalities by parental socioeconomic status, while the average effect was close to zero. However, as mentioned, the Swedish reform was a combination of later tracking and raising the minimum school-leaving age, preventing identification of the effect of tracking per se. Jones et al. (2014) analyzed the health outcomes of individuals who were born in the United Kingdom in 1958, 53 percent of whom lived in regions with a mixed-ability secondary school system, while the remaining 47 percent resided in regions with early tracking. Their results are mixed: in the late tracking regions, the worst-off types were doing better in terms of self-assessed health, but worse in terms of mental health, disability,

4

and chronic illness compared with those in the early-tracking regions.

The sign and size of the effects of late tracking on mortality are ex ante ambiguous and expected to differ by parental income. If tracking influences human capital accumulation, we expect the effects to persist over the life cycle and eventually to impact mortality. According to the tracking literature, peer effects and distance to the level of instruction in a classroom could matter for the outcomes of students. In addition, previous work has found that early tracking could lead to misclassification of students into a track that is either too demanding or too easy for them (Dustmann 2004). Given that theoretical predictions are ambiguous, our paper addresses the question which of these effects dominates.

We use unique administrative data on Finnish individuals who were in high school when the tracking reform took place in the 1970s. We can link these individuals while they were in their teens to the income of their parents, and observe mortality and hospitalizations until they reach their early fifties. For identification, we use a difference-in-differences approach that exploits the gradual implementation of the reform across six geographical regions between 1972 and 1977 to control for differences between regions that were stable across cohorts, and for differences by cohort that did not vary by region.

We find that later tracking led to a reduction in mortality among children from low-income families, but increased the mortality of children from high-income parents. Because the effects for children from low- and high-income families cancel out, we find no significant average effect of the reform. The net result is that delaying the educational tracking age in Finland in the 1970s reduced socioeconomic disparities in mortality which persisted throughout the life-cycle. We find no evidence of exposure to the new academic curriculum, which was introduced at the same time of the reform

5

but to which different cohorts were exposed at different ages. Our results are consistent with the conclusion by the OECD that early student selection has a negative impact on students assigned to lower tracks and exacerbates inequalities, without raising average performance (OECD 2012).

2 Tracking as a pathway between parental income and mortality during adulthood

In this section we outline a simple model, based on four determinants of academic achievement that considers three mechanisms through which tracking can influence human capital formation. Similar to the model of Cunha et al. (2006), the variation in early life human capital formation sets off a chain of events over the life course (see e.g. Campbell et al. 2014, for empirical evidence). Ultimately, this could influence mortality through life style choices and income (Mazumder 2012, Lindahl 2005).

The basic setup is that high- and low-performing students are in the same classroom with a one-size-fits-all curriculum and level of instruction until the tracking age. At that age, students are tracked into either an academic or a vocational track, depending on whether academic achievement is above or below a certain threshold value. The teacher who advises on the tracking can observe academic achievement in the form of his/her personal assessment of the student's capacities, as informed by the student's test scores. However, the tracking decision is based on imperfect information on a student's true potential, and therefore it is possible that the teacher takes the wrong tracking decision. The goal of the social planner is to set the tracking age such that it optimizes social welfare, which depends on the academic achievement of all students at the end of their educational careers.

6

2.1 Four determinants of academic achievement

We describe four major determinants of academic achievement in the literature to guide the interpretation of our empirical results. In our simple model, academic achievement at each age is a function of (i) ability, (ii) distance between a student's level of academic achievement and the level of instruction by the teacher, (iii) peer effects of classmates, and (iv) childhood circumstance.

Ability. Academic achievement throughout the educational career is determined exogenously by time-invariant ability, which cannot be observed directly (Card & Krueger 1996). Ability is transmitted intergenerationally and is therefore positively correlated with parental ability, which on average will have translated into higher parental socioeconomic status.

Distance to the level of instruction. Academic achievement suffers from greater absolute distance between the level of current academic achievement of a student and the level of instruction by the teacher (Duflo et al. 2011). The level of instruction itself depends on the average level of academic achievement of all students in a class. The negative effects of being taught at a level that is either too high or too low in the past persist at higher ages.

Classroom peer effects. Third, peer effects lead to spillovers: the student benefits from high academic achievement of his/her classmates (Sacerdote 2011). Academic achievement of a student is increasing in the achievement of his/her classmates. For example, high-performing students may explain the curriculum to their classmates, or less ambitious students convince classmates to skip classes.

Childhood circumstance. Given that a student at a younger age may not have revealed his/her true potential (Sacerdote 2011, Brunello & Checchi 2007, Malamud & Pop-Eleches 2011), childhood circumstance refers to the influence that parents can have on the academic achievement of their children. For example, high-income parents can afford private tutoring or test preparatory coaching for their children (Dustmann 2004, Hanushek et al. 2006, Bauer & Riphahn 2006, Brunello & Checchi 2007). Childhood circumstance becomes less important as a child gets older, in contrast to ability, the effect of which is time-invariant. This means that for a given level of time-invariant ability, we expect children from low-income families to do better in terms of academic achievement as they grow older, while we expect the opposite for children from high-income families. Childhood circumstance can thus be interpreted as a type of measurement error which leads to imperfect information about a student's true potential on the part of the teacher who gives the tracking advice.

2.2 Three mechanisms driving the heterogeneous effects of late tracking on academic achievement

Given the four determinants of academic achievement, we identify three theoretical mechanisms through which the tracking age could have potentially heterogeneous effects on human capital formation: (i) through peer effects, (ii) through the "distance" between the level of the student and the level of instruction, and (iii) through selection of students into a track that is either above or below their level, which we will refer to as mistracking.

Late tracking increases the duration of exposure to classmates with different academic levels. Compared to the selective system, students with high ability in mixed

classrooms suffer from the interaction with their low-ability peers, while students with low ability actually benefit from their high-ability classmates. Given that ability is positively correlated with parental socioeconomic status, peer effects in the mixed system would therefore favor students from low-income families and negatively affect the achievement of students from high-income families.

Late tracking increases the average distance to the level of instruction. Compared to the selective system, the level of instruction in the mixed system is higher than it would have been in the low track of the selective system, but lower than in the high track. This increases the average distance to the level of instruction. As a result, late tracking has adverse effects on both high- and low-ability students who are now taught at a level that is now either too low or too high (Betts et al. 2011). This mechanism suggests negative effects of late tracking for students from both low- and high-income families.

Late tracking decreases the probability of mistracking. Mistracking is more likely when students are tracked at a younger age because childhood circumstance obscures a student's true potential at a younger age. Mistracking is most likely to occur for students who are at the margin of being tracked in either the high or low track, in contrast to students at the top and bottom of the ability distribution. For these students, small differences in parental circumstance could shift academic achievement to the wrong side of the academic achievement threshold that determines track selection. Students from low-income families are more likely to be tracked too low ("undertracked") because academic achievement underestimates ability, while the opposite ("overtracking") is true for students from high-income families.

2.3 Theoretical predictions of the heterogeneous effects of late tracking by parental income.

Given the considerations in the previous subsections, we expect heterogeneous effects of late tracking by parental income. For those who would have been tracked appropriately in the selective system, late tracking negatively affects children from high-income families, while the predictions for children from poor families are mixed. Students with high ability, who are more likely to come from high-income families, may suffer from both the interaction with low-performing peers and the lower level of instruction in the mixed classrooms. For students with low ability, late tracking increases the distance to the level of instruction, but on the other hand it also leads to prolonged interaction with better-achieving peers.

Those who would have been mistracked in the selective system have more time to demonstrate their true potential in the mixed system because the role of childhood circumstance diminishes as they grow older. Because of the role of childhood circumstance, students from low-income families are more likely to be undertracked while the opposite is true for students from high-income families. Low-ability students, who would have been mistracked into the high track, now benefit from being instructed at a lower level, and mistracked high-ability students benefit from a higher level of instruction in the mixed classroom. The peer effects in mixed classrooms are advantageous for those who would have been undertracked in the selective system, but are worse for those who would have been overtracked.

There are theoretical reasons to expect heterogenous effects by parental income, but it is ex ante unclear whether and for whom late tracking is beneficial or detrimental to academic achievement. We estimate the total effect of these mechanisms taken together

in the context of the Finnish reform. This only tells us which of the theoretical mechanisms dominate the others, rather than giving us the contributions of each mechanism separately. This total effect, differentiated by parental income, is the policy-relevant effect: it forms a sufficient statistic for welfare comparison between the pre- and post-reform Finnish tracking policies.

The preceding discussion implies that we expect heterogeneous effects of the Finnish compulsory schooling reform, depending on childhood circumstance: children from low-income families may benefit from later tracking while their high-income counterparts may experience worse outcomes than would have been the case if they had been placed in a high-ability track early on. Indeed, previous work has established that late tracking reduced inequalities in education and income, and particularly benefited students with low-ability or from disadvantaged backgrounds (Pekkarinen et al. 2009, Kerr et al. 2013). Building on this evidence that the reform has had an effect on socioeconomic inequalities in human capital, we study whether these effects persisted over the life cycle and translated into longer and healthier lives. We then investigate whether any potential health and longevity gains for those from low-income parents are offset by losses for those from high-income parents.

3 Data on mortality and parental income in Finland

We use an 11 percent random sample of Finnish residents who were born between 1959 and 1965 and who were residing in Finland between 1987 and 2012, obtained from the Labor Market Data File and Census Records maintained by the Statistics Finland. Through a household identification number, we can link children to their parents and observe household income during childhood. Information on income is obtained from

Table I: Descriptive statistics

	Males			Females		
	All	Low parental income	High parental income	All	Low parental income	High parental income
Average hh inc. in 1970 and 1975	7,208 (3,761)	4,081 (1,593)	15,240 (9,897)	7,190 (3,728)	4,083 (1,592)	9,858 (2,872)
Proportion deceased	.0577	.0646	.0518	.0223	.0262	.0190
Observations	28,343	13,103	15,240	27,859	12,868	14,991

Notes: Sample means for men born between 1959 and 1965; standard deviations in parentheses. Top-coded household/parental income in 1975 prices. Mortality for the period between 1987 and 2012.

the tax register, and our main measure of family socioeconomic status is the inflation-adjusted average of household taxable incomes observed in 1970 and 1975 (i.e. during the childhood of the individual). Household income was top-coded at 17,705 in 1970 and at 17,800 in 1975 – both in 1975 prices.[1]

Our main outcome is mortality between 1987 and 2012. We observe the exact date of death for individuals who have died between 1987 and 2012, obtained from the cause-of-death registration. Given that we study the 1959-1965 cohorts, this implies we observe mortality in middle-age. Table I shows sample characteristics for males and females who were born between 1959 to 1965. Top-coded parental income was on average 76 percent higher in high-income households than in lower-income households.By the end of 2012, the year when the cohorts in the sample were around the age of fifty, 5.4 percent of individuals in the sample had died. Mortality among individuals who grew up in low-income households was 26 percent higher than in high-income households.

---

[1]Household income was measured during childhood, so we will use the term parental income interchangeably with household income.

4 The Finnish comprehensive school reform as a natural experiment to measure the impact of delayed tracking

Several key features of the Finnish comprehensive school reform make it particularly suitable for econometric evaluation of the effects of later tracking. Finnish children start primary school at the age of 7 (see figure 1) and 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 the 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 reform, tracking was postponed by five years until the age of 16. It meant that children were held together in mixed classes throughout the 9 years of comprehensive education.[2]

Two other features of the reform were that (i) compulsory schooling was extended to nine years, and (ii) a more academic curriculum was introduced (Pekkarinen et al. 2009). However, in practice almost all children already took at least nine years of schooling and therefore there was no effect of the reform on years of education (Pekkarinen 2008). Regarding the curriculum change, this co-occurred with the change of the tracking age and so the potential effect of the reform is a combination of later tracking and an adjusted curriculum.

The step-wise implementation meant that in each year between 1972 and 1977, one of six Finnish regions implemented the reform. Figure 2 illustrates the phasing of the implementation scheme, and shows that Helsinki was the last region to implement the reform. Our estimation strategy will use the fact that students in the early-

---

[2]The tracking reform did not affect between-school socioeconomic disparities, and there is ample interregional variation in socioeconomic conditions. Unfortunately we cannot observe whether certain parents moved between school districts.
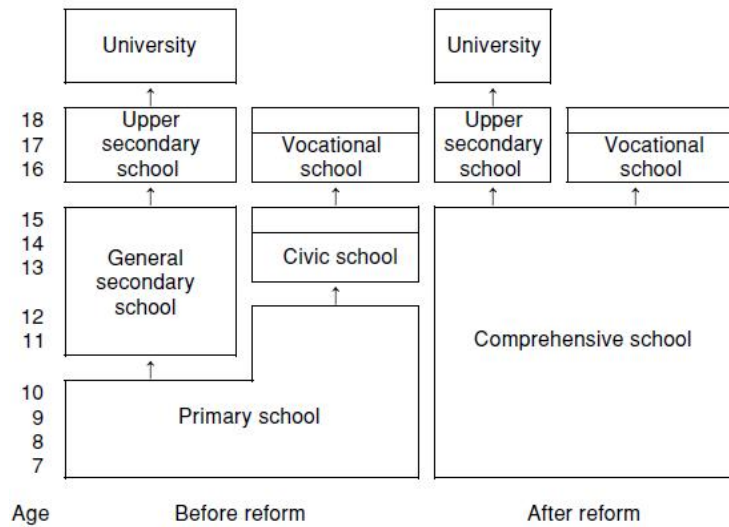
Figure 1: Finnish school system before and after the reform. Source: Pekkarinen (2009)

implementation regions were already taught under the new regime, while others in the same birth cohorts but in other implementation regions were still educated in the old system.

In the year of the reform, pupils who were beyond grade five (i.e., between 12 and 15 years old) continued in the old system, while students starting grades five and below were transferred to the new system (table II). To illustrate how the degree of exposure to the new curriculum differed by birth cohort in each region, consider the first implementation region, where the reform was implemented in 1972. Pupils born in 1960 or before were already beyond grade 5 by 1972 when their region implemented the reform. Hence, they were not exposed to the reform and belong to the control group. The 1961 birth cohort in the 1972 reform region, which 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 go through the full nine years of the comprehensive school system. The variation in the number
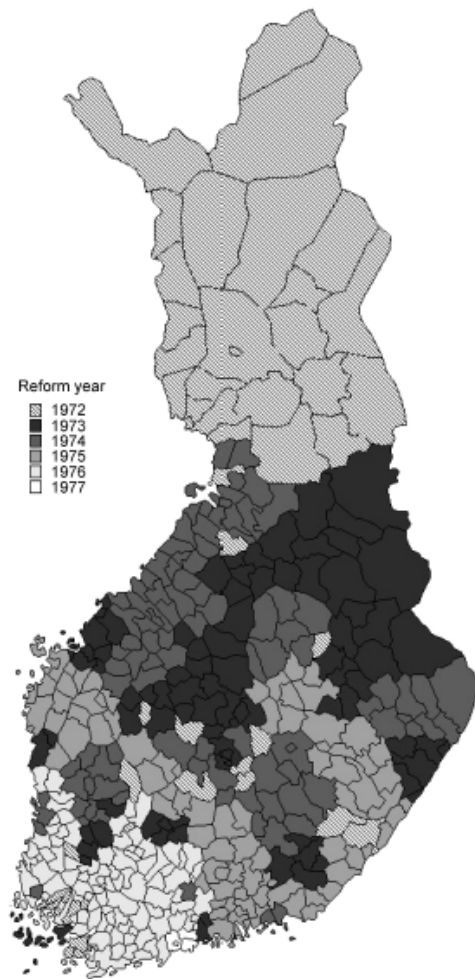
Figure 2: The year of the implementation of the reform varied between municipalities.
Source: Pekkarinen (2009)

Table II: Implementation of the reform

Birth cohort		Region					
		1972	1973	1974	1975	1976	1977
	1959	C	C	C	C	C	C
	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

Notes: 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.

of years of exposure among the first cohorts in the new system allows us to gauge the relative importance of the effect of tracking versus the effect of the new curriculum.

5 Estimation methods

We employ a difference-in-differences model to estimate the effect of the reform on mortality. We regress the outcome variable on a binary treatment variable t, six birth cohort dummies relating to the 1959 to 1965 birth cohorts in vector $\mathbf{c}$, and five reform region dummies in vector $\mathbf{r}$. The regression equation is:

$$y_{i,c,r} = \beta_0 + \beta_1 t_{c,r} + \boldsymbol{\beta}_1' \mathbf{c}_c + \boldsymbol{\beta}_2' \mathbf{r}_r + \varepsilon_{i,c,r} \tag{1}$$

Where the outcome variable y is an indicator of mortality between 1987 and 2012, for individual i in cohort c in region r. β_1 reflects the treatment effect of the reform.

Our identification strategy relies on the assumption that differences between treatment regions remain constant across the 1959 to 1965 birth cohorts. In other words, cohort effects are assumed to be the same for each of the six regions in the absence of

16

the reform. This is equivalent to the familiar common trends assumption in conventional difference-in-differences designs.[3]

We investigate the plausibility of the common trends assumption by testing the effect of a placebo treatment. The placebo-treated group is the cohort that was born one year before the reform cohort in each region as shown in Panel A of Table III. The idea is that if the cohorts who were one year too old to be affected by the reform in a given region produce a significant treatment effect, this is evidence that the common trends assumption is not satisfied (Autor 2003, Angrist & Pischke 2008). In our test, we exclude the treated cohorts in each region.

The lower panel of Table III illustrates our strategy to gauge the relative importance of the tracking and changed curriculum aspects of the reform. Our reasoning is that all treated cells in Table II are exposed to later tracking: e.g., the first affected group (the first diagonal with "T" in Table II) was 11 at the time of the reform, and the tracking age was postponed to age 16. Therefore, their exposure to the new curriculum was 5 years. For the triangle below the main diagonal in Table II), their tracking age was also extended to age 16, but their exposure to the new curriculum was 6, 7, 8 or even 9 years because the new curriculum was implemented in the year of the reform, also for younger cohorts. Hence, one would expect that if tracking were the main driver of any potential reform effects, one would observe this already on the first diagonal; while if the change in curriculum were mainly driving the effects of the reform, then one would expect the effects to be larger among the lower triangle of the Table. To gauge the relative importance of these two effects, we therefore present additional estimates based on an extended version of estimation equation 1, now including a dummy representing

---

[3]In contrast to Pekkarinen et al. (2009), we exclude the 1966 birth cohort because all individuals in this cohort are treated.

Table III: Placebo test and intensity

A. Placebo		Region					
		1972	1973	1974	1975	1976	1977
	1959	0	0	0	0	0	0
	1960	1	0	0	0	0	0
	1961		1	0	0	0	0
Birth cohort	1962			1	0	0	0
	1963				1	0	0
	1964					1	0
	1965						1

B. Intensity		Region					
		1972	1973	1974	1975	1976	1977
	1959	0	0	0	0	0	0
	1960	0	0	0	0	0	0
	1961	0	0	0	0	0	0
Birth cohort	1962	1	0	0	0	0	0
	1963	1	1	0	0	0	0
	1964	1	1	1	0	0	0
	1965	1	1	1	1	0	0

Notes: Panel A shows the values of the placebo dummy variable, empty cells are missing. Panel B shows the values of the intensity dummy. 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.

the treatment group below the main diagonal, as illustrated in the bottom panel of Table III.

On the basis of our theoretical framework and the results of Pekkarinen (2008) and Pekkarinen et al. (2009), we anticipate the reform to have heterogeneous effects by gender and parental income. Therefore, we estimate equation 1 separately for men and women who grew up in poor and rich families. Poor families are those below the median of family income, and rich families are defined as those above median family income.

We use OLS estimation, as well as Cox proportional hazard models. The Cox proportional hazard technique is an estimation method that relies on the assumption of common trends in the index function because of the nonlinear transformation of the link function (Lechner 2011, Norton et al. 2004, Puhani 2012). The advantage of the Cox proportional hazard model is that it exploits the duration dimension of the data, using more variation in the data than just the binary indicator of mortality. The survival data are left-truncated because we observe survival from 1987 onwards, and right-censored at January 1, 2013. We cluster standard errors at the regional level, which is the level of variation in the treatment variable. This leads to higher threshold values for the t-statistic for each given significance level (Angrist & Pischke 2008).

6 Results

The first row of panel A of table IV shows the association between mortality and a dummy for high parental income among men (column 1) and women (column 4) who were not affected by the reform: the cohort-region cells that are denoted with a C in table II. On average, high parental income is associated with a 1.6 percentage points

(or 23 percent) reduction in mortality among men, and a 0.7 percentage points (or 26 percent) reduction in mortality among women. Hence, while mortality is clearly lower among women, the association between parental income and mortality is similar across gender.

The estimated average treatment effects for men and women with low-income parents are presented in columns 2 and 5, respectively. While for women the estimates are insignificant, the estimates for men are significant at the 10 percent significance level, suggesting that the reform has improved survival for men with low-income parents.[4]

Column (3) and column (6) show the effect of the reform on men and women from high-income families. Again, for women we do not find any effect of the reform, while the estimate of the effect of the reform on men from high-income parents suggests that it has *reduced* survival among this group, and significantly so at the 5 percent level. The coefficients for the high and low income groups differ significantly from each other (p=0.009).[5] Panel B of Table IV repeats the analyses using a Cox proportional hazard model. Columns 1 and 4 show that the estimated probability of dying for men (women) who were raised in high-income households is 24 (26) percent lower compared to those from low-income families throughout the life cycle, pretty much in line with the linear model. As in the linear specification, we do not find evidence that the reform had an impact on mortality among women. In contrast, column 2 shows evidence that the reform has substantially improved longevity for men who were raised in low-income

---

[4]The sample is split at the median, which is determined by an augmented regression that is similar to equation 1 but includes linear interactions between all covariates and the logarithm of parental income. The predicted effect of the reform in this model is positive for those above the median, and negative below that threshold. The p-value of the interaction term between the reform and log parental income is .114.

[5]The effects of opposite sign for men from high- and low-income families cancel out on aggregate. The estimated effect for the pooled sample of men is .0011 (.0044) and for the pooled sample of women it is .0014 (.0019).

Table IV: The effect of the reform on mortality: ordinary least squares and Cox proportional hazard estimation

	Males			Females		
	(1) Before reform	(2) Low income	(3) High income	(4) Before reform	(5) Low income	(6) High income
A. Linear regression						
Constant	.0702*** (.0027)			.0282*** (.0017)		
High parental income	-.0161*** (.0035)			-.0074*** (.0032)		
Reform		-.0155* (.0065)	.0104** (.0034)		-.0015 (.0021)	-.0016 (.0027)
Region and cohort		✓	✓		✓	✓
Observations	19,149	13,103	15,240	18,869	12,868	14,991
B. Cox proportional hazard						
High parental income	.7640*** (.0482)			.7369*** (.0484)		
Reform		.7859** (.0895)	1.2672*** (.0820)		.9259 (.0780)	.9096 (.1447)
Region and cohort		✓	✓		✓	✓
Observations	19,149	13,103	15,240	18,869	12,868	14,991

Notes: The sample consists of men and women who were born between 1959 and 1965. The coefficients of the cohort and region dummies 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. Low parental income is the reference category.
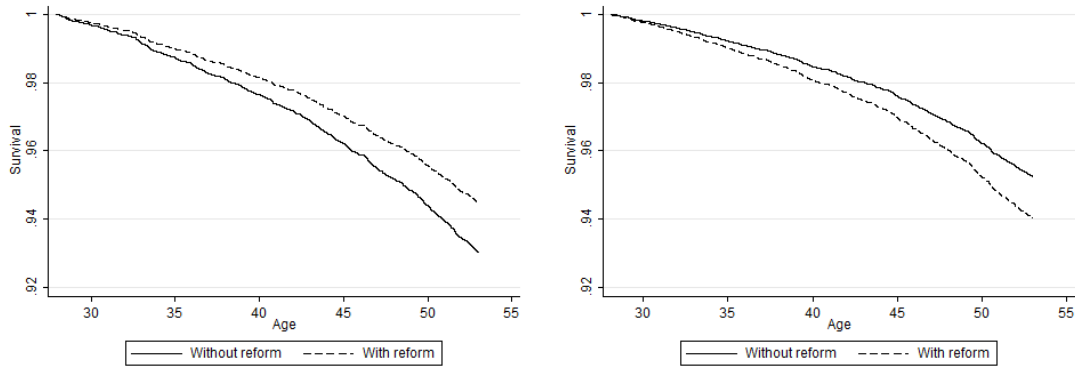
(a) Survival after age 28 for men who grew up in low-income households

(b) Survival after age 28 for men who grew up in high-income households

Figure 3: Effect on survival, split by parental income.

families (the point estimate suggests a 21.4 percent decrease in the probability of death from the baseline), while column (3) shows that the reform increased mortality by almost 27 percent for men who were raised in high-income families. This confirms the OLS result that the reform has thus narrowed mortality differentials among men, and that the magnitude is roughly 1 percentage point in absolute terms and roughly 25 percent relative to the baseline mortality of 5 percent..

Figures 3a and 3b visualize the predicted survival curves for men, both before and after the reform. The figure starts at age 28, the age of the 1965 birth cohort in 1987 which is the year in which we first observe mortality. The survival curve of men from low-income families before the reform (the solid line in figure 3a) lies well below the survival curve of those from high-income parents before the reform (the solid line in figure 3b). The dashed lines in both figures plot the (predicted) survival curves after the reform. After the reform, the difference between men from lower income families and men from higher income families is no longer significant (results not shown).

The findings from the duration model confirm our finding in panel A that the effect

22

of the reform varies by parental income for men. In line with the results of Pekkarinen et al. (2009), the effect of the reform on women is statistically insignificant. For men, the evidence consistently shows that the longevity gains of boys from lower-income families came at the expense of a reduced longevity among boys from richer families. Next, we subject the results for men to a number of tests.

Common trends assumption The coefficients of the placebo dummy in columns (1) and (2) of both panels of table V should not be different from zero if cohort trends in the outcome variable did not differ by region. The two coefficients in panel A are close to zero and have p-values of .33 and .85, which means we cannot reject the common trends assumption. The p-values in panel B are .21 and .98. When we include the placebo dummy alongside the treatment variable in the same model we obtain the same results (available upon request). This supports the notion that our findings are not driven by factors other than the reform.

Tracking versus curriculum change Table VI presents results from augmented regression models including the "intensity" dummy that indicates whether an individual is part of the lower triangle from panel B of table III and therefore was exposed to the curriculum change from a younger age than the first reform cohorts.

The coefficients of the intensity dummies are never significant in any of the models, and therefore we cannot reject the hypothesis that more years of exposure to the new curriculum had no effect.

Table V: The effect of the reform on mortality: placebo tests

	(1) Low parental income sample	(2) High parental income sample
A. *Linear regression model*		
Placebo	.0063	.0008
	(.0058)	(.0041)
Region and cohort dummies	✓	✓
Observations	8,136	11,013
B. *Cox proportional hazard model*		
Placebo	1.1152	.9972
	(.0959)	(.0982)
Region and cohort dummies	✓	✓
Observations	8,136	11,013

Notes: The sample consists of men who were born between 1959 and 1965 and who were not exposed to the reform. The coefficients of the cohort and region dummies 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. Low parental income is the reference category.

7 Conclusion

To the best of our knowledge, our paper is the first to study the effect of late tracking on socioeconomic inequalities in mortality. Exploiting regional variation in the moment of tracking in Finland, our difference-in-differences framework accounts 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 that delaying the tracking age generally reduced the gradient between socioeconomic circumstance during childhood and mortality later in life among men. Yet, strikingly, the longevity gains for children from poorer families have come at the cost of longevity losses to children from affluent families, and so there are both winners and

Table VI: The effect of the reform on mortality: tracking versus intensity

	(1) Low parental income sample	(2) High parental income sample
A. Linear regression model		
All reform cohorts	-.0178**	.0087**
	(.0057)	(.0033)
Later reform cohorts	.0091	.0054
	(.0092)	(.0068)
Region and cohort dummies	✓	✓
Observations	13,103	15,240
B. Cox proportional hazard model		
All reform cohorts	.7332***	1.2158***
	(.0871)	(.0784)
Later reform cohorts	1.2211	1.1344
	(.2070)	(.1779)
Observations	13,103	15,240

Notes: The sample consists of men who were born between 1959 and 1965. The coefficients of the cohort and region dummies 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.

losers of late tracking.

We could not offer conclusive evidence about the reasons why the reform differentially impacted mortality. We have investigated potential mechanisms of the effect of late tracking on mortality using hospitalization and medicine purchase registers. In particular, we investigated three outcomes that are prevalent in middle-age (WHO, 2004) : accidents, hypertension medication and the purchase of antidepressants. While the point estimates were in line with our main estimates for mortality (i.e., lower accidents and use of medication among the lower income groups, and more accidents and medication use among the higher income groups), these estimates lack statistical power, most probably because of the low prevalence of serious health conditions in middle age.

Nonetheless, building upon the literature and our own suggestive evidence, a possible interpretation emerges. The findings from Pekkarinen et al. (2009) and Kerr et al. (2013) confirmed that the fall in intergenerational income inequality derived from the improved test scores of lower income boys. Hence, reduced income inequality is one possible mechanism for the reduced inequality in mortality by income. Moreover, and more generally, the tracking literature suggests that tracking has heterogeneous effects on human capital through exposure to a different set of peers, different knowledge obtained in school, and a lower probability of mistracking. These mechanisms may have eventually contributed to the heterogeneous effects on mortality that we found.

While the reduction of inequalities in various outcomes with respect to parental socioeconomic status will primarily be considered a success, and while policymakers are possibly willing to accept some reduction in income among the better-off, it is plausible that the size of the penalty that is paid in terms mortality was unintended. Our paper therefore tells a cautionary tale that while late tracking may indeed be successful

in reducing socioeconomic disparities, this may come at a price: deteriorated health outcomes among the better-off. Various puzzles remain. It is not clear, for instance, whether a reduction in within-school selection because of later tracking is followed by increased between-school selection at the primary school starting age, and why the reform had no effect on women. Future follow-up may reveal whether the reform effects among men are maintained at older ages.

Our results show that, on average, tracking at younger ages did not lead to any significant changes in mortality, partly reflecting the mortality redistribution that occurred as a result of the reform. Because tracking at younger ages is less likely to benefit the worse-off, one may argue that later tracking is a fairer and just as efficient policy. Yet, a full welfare analysis would be required in order to establish the long-term benefits and costs of this policy.

References

Ammermueller, A. (2013), 'Institutional features of schooling systems and educational inequality: Cross-country evidence from pirls and pisa', *German Economic Review* **14**(2), 190–213.

Angrist, J. D. & Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Autor, D. H. (2003), 'Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing', *Journal of Labor Economics* **21**(1), 1–42.

Bauer, P. & Riphahn, R. T. (2006), 'Timing of school tracking as a determinant of intergenerational transmission of education', *Economics Letters* **91**(1), 90–97.

Betts, J. et al. (2011), 'The economics of tracking in education', *Handbook of the Economics of Education* **3**(4), 341–381.

Brunello, G. & Checchi, D. (2007), 'Does school tracking affect equality of opportunity? new international evidence', *Economic Policy* **22**(52), 782–861.

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

Card, D. & Krueger, A. B. (1996), 'School resources and student outcomes: An overview of the literature and new evidence from north and south carolina', *The Journal of Economic Perspectives* **10**(4), 31–50.

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

Cesarini, D., Lindqvist, E., Östling, R. & Wallace, B. (2016), 'Wealth, health, and child development: Evidence from administrative data on swedish lottery players', *Quarterly Journal of Economics* **131**(2), 687–738.

Clark, D. & Royer, H. (2013), 'The effect of education on adult mortality and health: Evidence from britain', *The American Economic Review* **103**(6), 2087–2120.

Cunha, F., Heckman, J. J., Lochner, L. & 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. & Kremera, M. (2011), 'Peer effects, teacher incentives, and the

impact of tracking: Evidence from a randomized evaluation in kenya', *The American Economic Review* **101**(5), 1739–1774.

Dustmann, C. (2004), 'Parental background, secondary school track choice, and wages', *Oxford Economic Papers* **56**(2), 209–230.

Grossman, M. (2015), 'The relationship between health and schooling: What's new?', *Nordic Journal of Health Economics* **3**(1), pp–7.

Hanushek, E. A. et al. (2006), 'Does educational tracking affect performance and inequality? differences-in-differences evidence across countries*', *The Economic Journal* **116**(510), C63–C76.

Jones, A. M., Roemer, J. E. & Dias, P. R. (2014), 'Equalising opportunities in health through educational policy', *Social Choice and Welfare* **43**(3), 521–545.

Kerr, S. P., Pekkarinen, T. & Uusitalo, R. (2013), 'School tracking and development of cognitive skills', *Journal of Labor Economics* **31**(3), 577–602.

Lechner, M. (2011), 'The estimation of causal effects by difference-in-difference methods', *Foundations and Trends in Econometrics* **4**(3), 165–224.

Lindahl, M. (2005), 'Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income', *Journal of Human resources* **40**(1), 144–168.

Malamud, O. & Pop-Eleches, C. (2011), 'School tracking and access to higher education among disadvantaged groups', *Journal of Public Economics* **95**(11), 1538–1549.

Mazumder, B. (2012), 'The effects of education on health and mortality', *Nordic Economic Policy Review: Economics of Education* pp. 261–301.

Meghir, C., Palme, M. & Simeonova, E. (2013), Education, cognition and health: Evidence from a social experiment, Working Paper 19002, National Bureau of Economic Research.

Norton, E. C., Wang, H. & Ai, C. (2004), 'Computing interaction effects and standard errors in logit and probit models', *Stata Journal* **4**, 154–167.

OECD (2012), *Equity and Quality in Education-Supporting Disadvantaged Students and Schools*, OECD Publishing.

Palme, M. & Sandgren, S. (2008), 'Parental income, lifetime income, and mortality', *Journal of the European Economic Association* **6**(4), 890–911.

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

Puhani, P. A. (2012), 'The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models', *Economics Letters* **115**(1), 85–87.

Sacerdote, B. (2011), 'Peer effects in education: How might they work, how big are they and how much do we know thus far?', *Handbook of the Economics of Education* **3**, 249–277.

Van Kippersluis, H., O'Donnell, O. & Van Doorslaer, E. (2011), 'Long-run returns to education does schooling lead to an extended old age?', *Journal of human resources* **46**(4), 695–721.

Woessmann, L. (2016), 'The importance of school systems: Evidence from international differences in student achievement', *The Journal of Economic Perspectives* **30**(3), 3–31.