

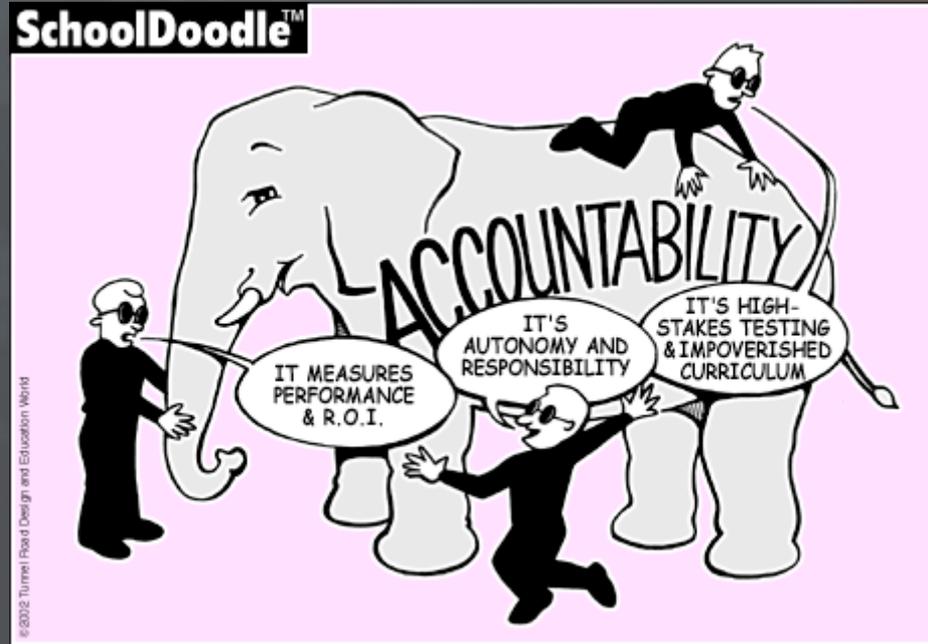
Effects of School Accountability on Social Inequality in Educational Attainment: A Comparative Reform Study of Europe

Herman van de Werfhorst
University of Amsterdam

(partly based on work with Jennifer L. Jennings, Princeton University)

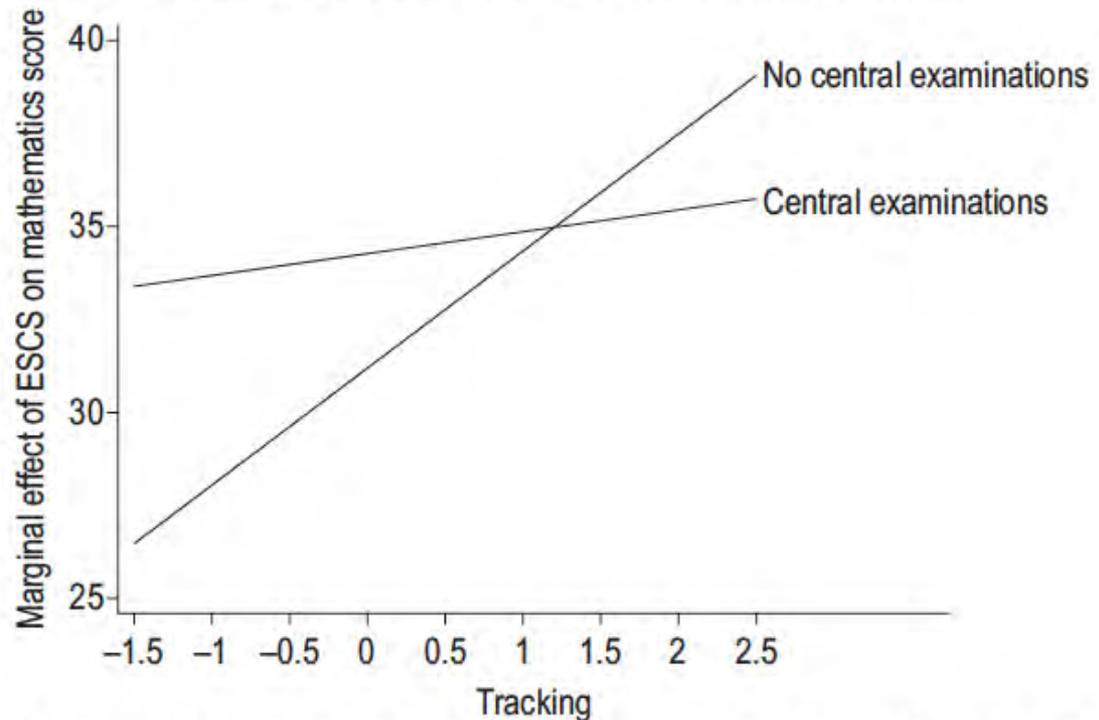
The many sides of school accountability

- Standardized tests?
- External school evaluation?
- Binding teacher recommendations?
- How do these forms of standardization relate to social inequality in schooling outcomes?



Earlier we found that centralized examinations were associated to reduced inequalities in early-tracking systems

Figure 2. Marginal effects graph of the three-way interaction (table 3)



Note: Based on model 4, table 3. There is a significant difference between the slopes for tracking in the absence of central examinations and presence of central examinations. ESCS refers to the PISA measure of socio-economic status.

Bol, Thijs, Witschge, Jacqueline, Van de Werfhorst, Herman G. and Dronkers, Jaap. (2014). Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries. *Social Forces*, 92(4): 1545-1572.

The Dutch reform with the final “cito” test is supporting the same hypothesis: standardization reduces inequalities by social background.

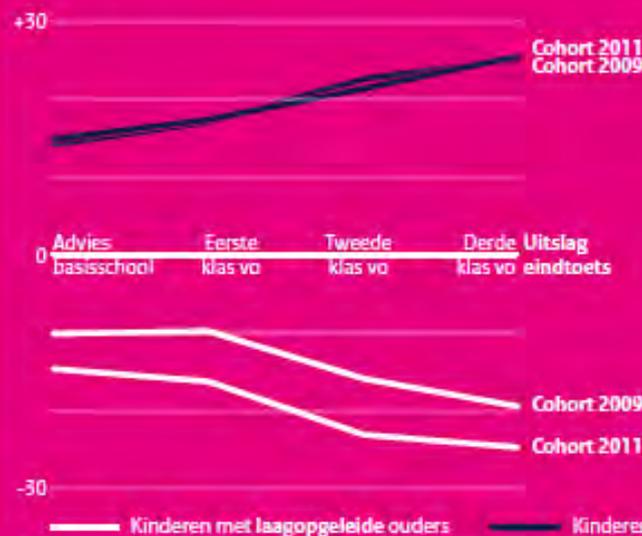
2

De verschillen nemen toe

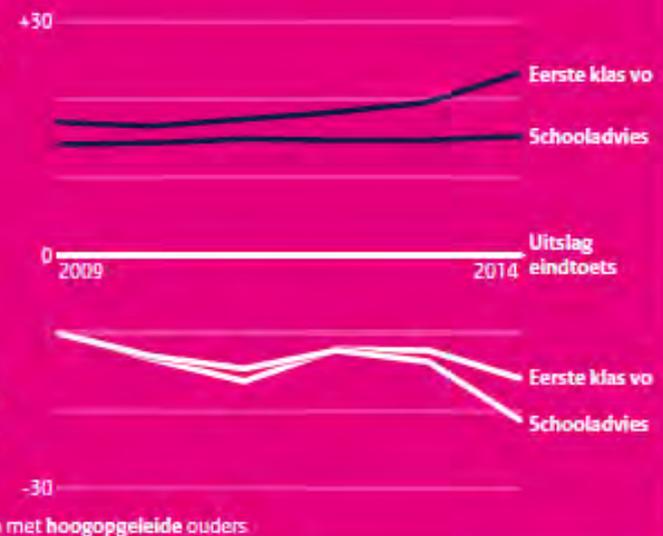
Kinderen met hoogopgeleide ouders krijgen van hun basisschool vaker een advies voor het voortgezet onderwijs dat hoger ligt dan de uitslag van de eindtoets. Bij kinderen met laagopgeleide ouders is dat andersom. Tijdens de onderbouw van het voortgezet onderwijs nemen deze verschillen toe.

De schooladviezen aan kinderen van laagopgeleide ouders dalen de afgelopen jaren. Kinderen van hoogopgeleide ouders beginnen op een hoger niveau dan in eerdere jaren. Ook daardoor worden de verschillen groter.

Gemiddelde afwijking van de eindtoets-uitslag in procenten ...
... per cohort ...



... en per kalenderjaar





So: does standardization reduce inequalities in education?

Standardization and accountability

A number of American studies on the effects of accountability on student outcomes

- No Child Left Behind Act: States used student test scores to evaluate schools (with major potential consequences).
- State-level studies.
- Summary Hout & Elliott 2011 (NRC report): small positive effects of average performance (mostly in elementary school, not in middle school).

But some studies point to larger inequalities as a consequence of NCLB

- Educational triage: neglect the low-achievers.
- Teaching to the test: neglect of 'untested outcomes' such as long-term educational attainment or untested math.
- Recent pleas for **Education Inspectorates** rather than test-based accountability to deal with the problems (e.g. Helen Ladd, JPAM 2017). Working with schools, rather than against them.
- However, in the US there are no inspectorates, and little is known whether inspectorates are effective in other countries.

Helen Ladd (Duke University) in JPAM 2017

- “An alternative to the test-based accountability of NCLB would be to follow the lead of many other countries (...) by sending review teams to individual schools and districts on a periodic basis to review their quality and effectiveness”.
- Inspectorates are “specifically intended to be more formative, with the inspection process itself driving and supporting school improvement”.
- Inspectorates work *with* schools, rather than *against* them.

A look at Europe

- In Europe, we have witnessed reforms towards external school evaluations (i.e. inspectorates) and standardized tests.
- Research question:
 - Does the introduction of external school evaluation policies help in raising average educational attainment and reducing inequalities by social background?
 - And the introduction of standardized tests? And in combination?

Intended and unintended effects of external school evaluation

- **Intended**: improving performance, reducing inequities.
- **Unintended**: strategic action of schools, window-dressing, excluding weak students from the assessment. May reduce educational outcomes and increase inequity
- Arguably, it may be 'easier' for schools to act strategically this in 'holistic' school evaluations than on the basis of (individual) standardized tests.

Comparative reform study

- Socioeconomic inequalities in educational attainment.
- European Social Surveys rounds 1-7 (2002-2014)
- Born 1945-1989
- Time-varying indicators of reforms available for 20 countries. Reforms: **measures the direction of change (0/1)**, not the content of the reforms. Within-country comparison.
- Educational reforms merged to individuals when they were 11 years old.

Braga, Michela, Daniele Checchi, and Elena Meschi. 2013. "Educational Policies in a Long-Run Perspective." *Economic Policy* 28(73):45–100.

Checchi, Daniele, Van De Werfhorst, Herman G. 2017. "Policies, Skills and Earnings: How Educational Inequality Affects Earnings Inequality." *Socio-Economic Review*, advance access.

Difference in difference model

Before and after reform. Add country-specific inequality trend.

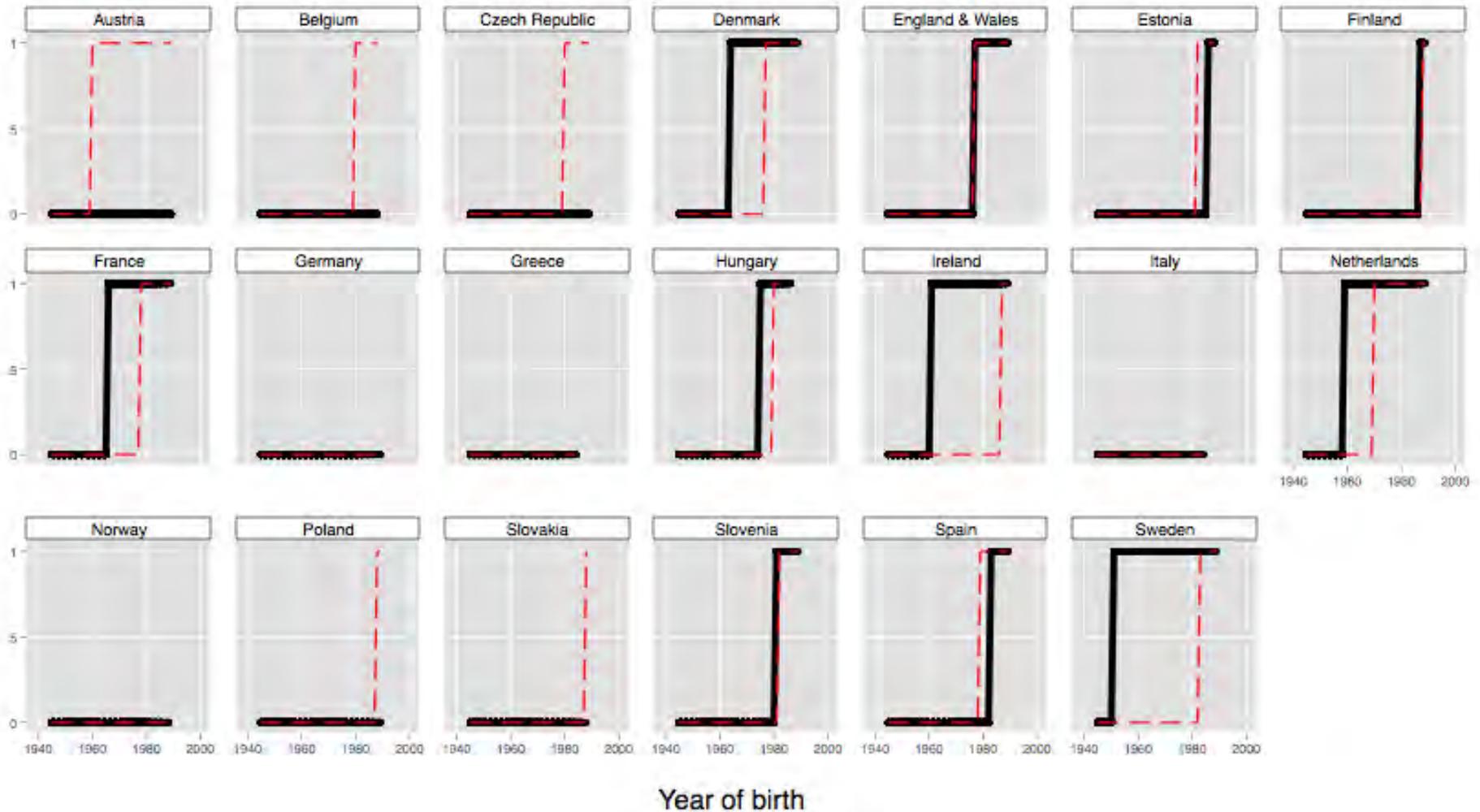
Given the historical reform data, we focus on long-term outcomes: Completion of upper secondary education, and highest qualification level achieved.

Linear models on educational attainment (less than lower secondary, lower secondary, upper secondary, tertiary) and Linear probability models on dichotomous outcomes.

$$\begin{aligned} Y_{itc} = & \alpha + \beta_0 BY_{itc} + \beta_1 Peduc_{itc} + \beta_2 Peduc_{itc} \cdot BY_{itc} + \beta_3 Peduc_{itc} \cdot CD_c \\ & + \beta_4 Peduc_{itc} \cdot BY_{itc} \cdot CD_c + \beta_5 BY_{itc} + \beta_6 CD_c \cdot BY_{itc} + \sum_{c=c}^{c-1} \gamma_c CD_c \\ & + \delta_1 Test_{tc} + \delta_2 SEV_{tc} + \delta_3 Test_{tc} \cdot Peduc_{itc} + \delta_4 SEV_{tc} \\ & \cdot Peduc_{itc} + \delta_5 Test_{tc} \cdot SEV_{tc} + \delta_6 Test_{tc} \cdot SEV_{tc} \cdot Peduc_{itc} + \varepsilon_{itc} \end{aligned}$$

Reforms in Europe

— Standardized test
- - External school evaluation



EXAMPLES

- Belgium/Flanders
A royal Decree was issued in 1991 for general school inspections, bridging several educational umbrellas.
- Estonia
In 1993, the Ministry of Culture and Education was created, and its area of government included the State Board of Education as a government institution, carrying out state supervision of education institutions and checking the fulfilment of education related legal acts and national education standards.
- Finland
Evaluation was stipulated by law in 1999; this obligates schools and educational institutions to perform self-evaluation. Also, a national system for evaluating learning outcomes was established.
- France
A unified system of inspectorates was created by law in 1989.
- Netherlands
1981: the law of foundational education; Inspectorate placed in between the ministry and the field.

Causal?

- The effects of reforms are causal to the extent that all unobserved contextual factors that may affect inequality do so through main inequality differences or country-specific inequality trends.
- E.g. broader national cultures of accountability may falsify this assumption but only to the extent that these do not translate in the country-specific inequality trend.
- And we test effects of alternative accountability measures too.
- And we have a placebo test.

Fit statistics tell us that **only external evaluation** is associated to the slope of parents' education

Model	Men					Women				
	BIC	Log Likelihood	Chi-square test (df)	sig.	against model	BIC	Log Likelihood	Chi-square test (df)	sig.	against model
<i>Without interactions with parents' education</i>										
1 Only school evaluation	-691827.70	-23103.77				-775292.49	-27124.98			
2 Only standardized test	-691825.68	-23104.78				-775264.37	-27139.04			
3 Both institutional variables	-691823.16	-23100.48	6.57 (1)	*	1	-775294.97	-27118.14	13.69 (1)	***	1
			8.59 (1)	**	2			41.81 (1)	***	2
4 Both institutional variables interacted	-691813.13	-23099.94	1.08 (1)	n.s.	3	-775284.57	-27117.73	0.81 (1)	n.s.	3
<i>With interactions with parents' education</i>										
5 Only school evaluation	-691821.58	-23090.17	27.20 (3)	***	1	-775289.45	-27109.68	30.60 (3)	***	1
6 Only standardized test	-691798.37	-23101.77	6.01 (3)	n.s.	2	-775233.22	-27137.80	2.49 (3)	n.s.	2
7 Both institutional variables	-691786.91	-23085.29	30.39 (6)	***	3	-775264.65	-27099.65	36.97 (6)	***	3
			9.76 (4)	*	5			20.06 (4)	***	5
			32.97 (4)	***	6			76.29 (4)	***	6
Both institutional variables, only evaluation										
8 interacted with parents' education	-691816.00	-23087.41	4.23 (3)	n.s.	7	-775293.10	-27102.25	5.20 (3)	n.s.	7
			5.52 (1)	*	5			14.86 (1)	***	5
9 Three-way interactions	-691745.77	-23083.64	32.60 (9)	***	4	-775222.93	-27098.08	39.29 (9)	***	4
			3.29 (4)	n.s.	7			3.13 (4)	n.s.	7

Note: these fit statistics are calculated on a linear probability model predicting completion of upper secondary education

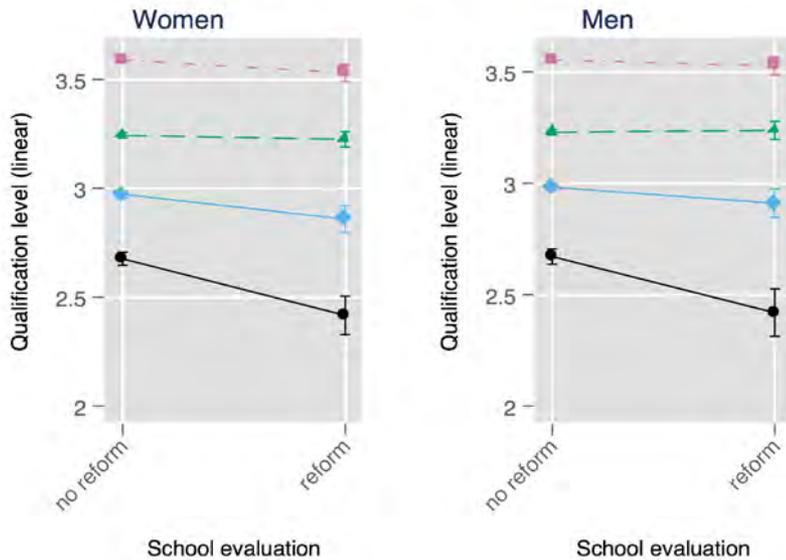
~ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Main results: Reform towards external evaluation is associated to lower attainments of children of low-educated backgrounds

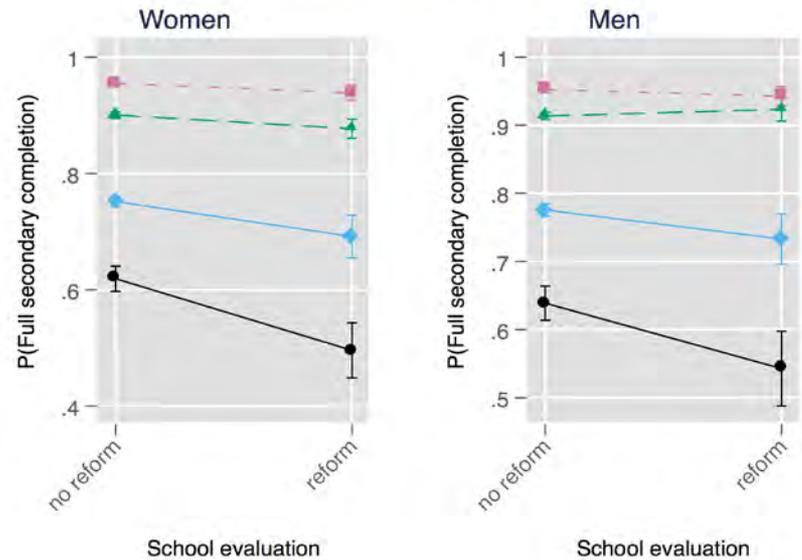
	Highest level of education		Completion upper secondary education		Quasi-Placebo: Bachelor's degree (conditional on upper secondary completion)	
	Men	Women	Men	Women	Men	Women
	Reform: Standardized test	-0.028 (0.018)	0.051** (0.019)	-0.018* (0.009)	0.028** (0.009)	0.008 (0.012)
Reform: External school evaluation	-0.252*** (0.055)	-0.260*** (0.046)	-0.096*** (0.027)	-0.123*** (0.023)	0.011 (0.035)	0.034 (0.030)
Parental education (relative to less than lower secondary)						
Lower secondary	0.104 (0.092)	0.167* (0.074)	0.026 (0.053)	0.045 (0.045)	0.012 (0.044)	0.062 (0.039)
Upper secondary	0.364*** (0.095)	0.342*** (0.077)	0.230*** (0.055)	0.223*** (0.050)	0.063 (0.042)	0.049 (0.033)
Tertiary	0.751*** (0.088)	0.785*** (0.080)	0.292*** (0.053)	0.299*** (0.047)	0.389*** (0.060)	0.421*** (0.041)
Parental education * External school evaluation reform						
Lower secondary	0.180** (0.064)	0.146** (0.054)	0.054~ (0.031)	0.063* (0.029)	0.003 (0.045)	-0.049 (0.037)
Upper secondary	0.260*** (0.060)	0.242*** (0.048)	0.106*** (0.028)	0.100*** (0.024)	-0.007 (0.040)	-0.011 (0.032)
Tertiary	0.224*** (0.057)	0.201*** (0.050)	0.086** (0.028)	0.107*** (0.024)	-0.015 (0.038)	-0.063~ (0.035)

Predicted outcomes

Qualification level attained



Full secondary completion



Parents' highest educational attainment

- incomplete secondary
- ◆— lower secondary
- ▲— upper secondary
- -■ - - tertiary

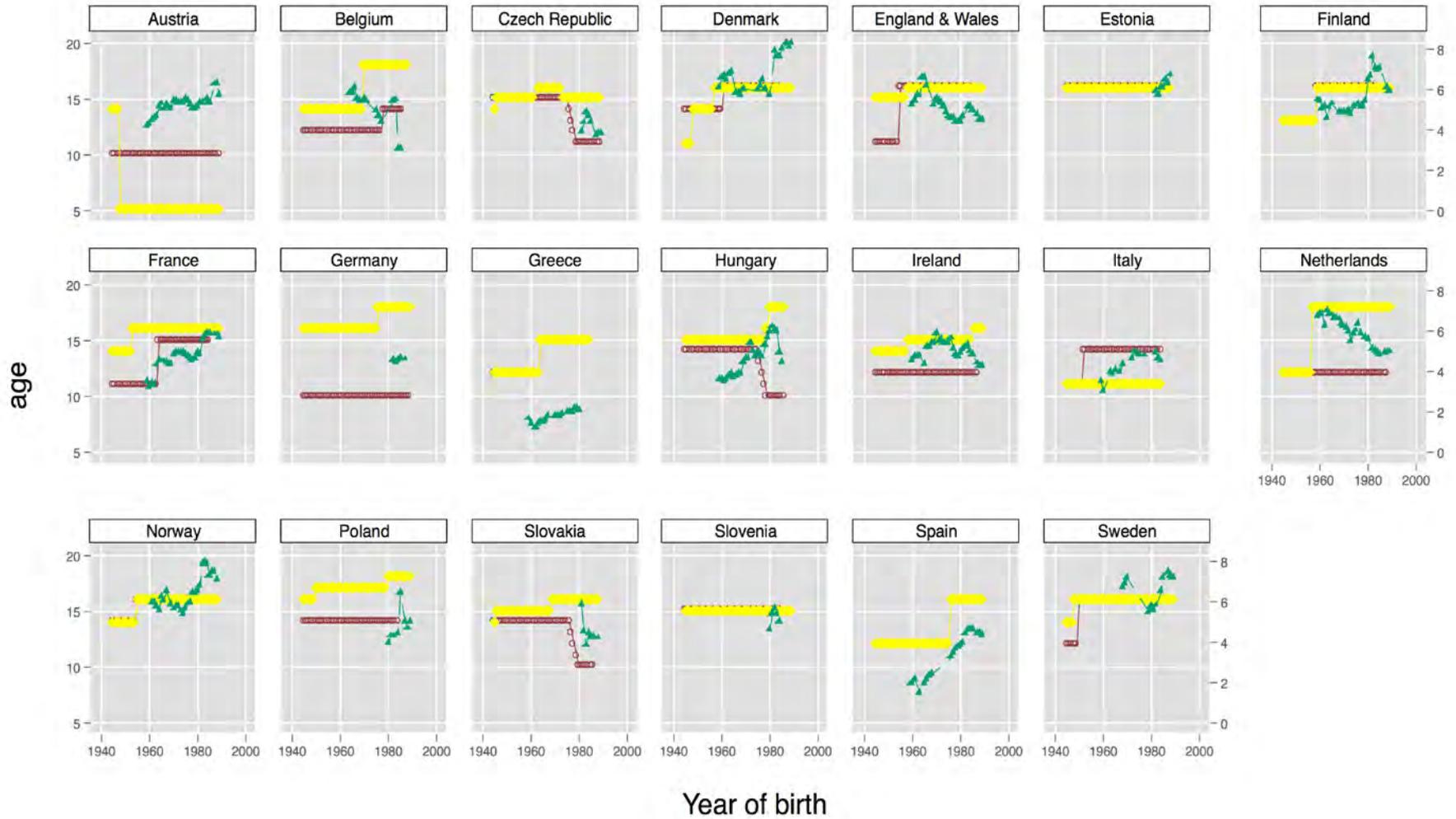
Testing alternative market-oriented reforms: school autonomy and teacher autonomy

Effect of external school evaluation is still significant and of similar size. School autonomy has an independent effect leading to larger inequalities (through the lowest-SES children).

Why? We test three explanations

- **Compulsory school leaving age?** External evaluation may be particularly influential if early exits are made possible.
- **Early tracking?** External evaluation may be particularly influential if lower tracks are available to divert low-SES students to.
- **Education expenditure?** Reforms may be more influential in societies with restricted government budgets for education.

- Tracking age
- ◆— Minimum compulsory school leaving age
- ▲— Public education expenditure as % of GDP (right Y-axis)

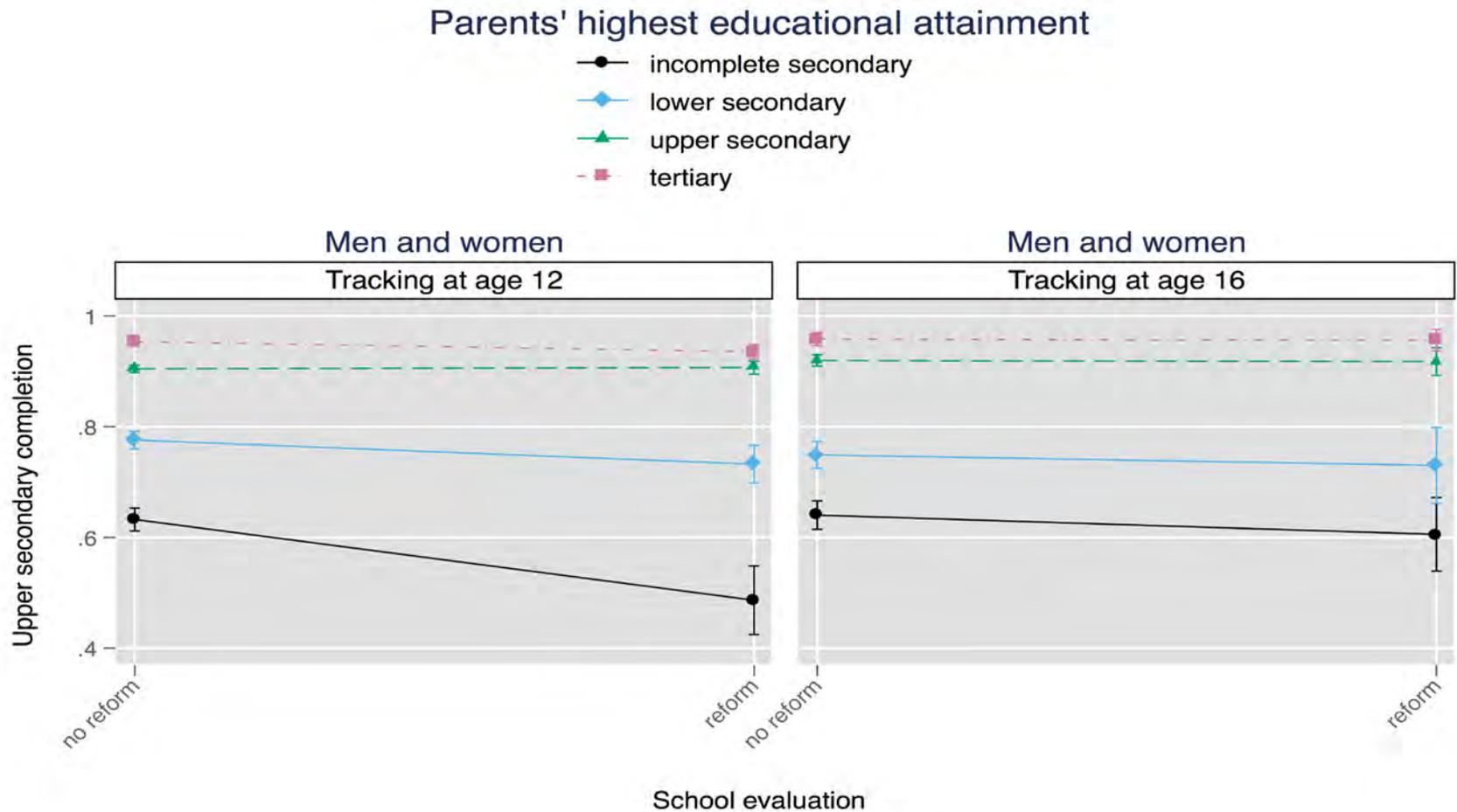


Fit statistics: only tracking age is associated to the evaluation effect on socioeconomic inequality

	Model 1	Model 2	
additional contextual variable:	Log Likelihood without interaction term	Log Likelihood with interaction term	Chi-Square difference sig.
Tracking age	-45703.701	-45694.111	19.18 (7) **
Compulsory schooling age	-50778.714	-50773.803	9.82 (7) n.s.
Public education expenditure as % of GDP	-18446.739	-18441.859	9.76 (7) n.s.

Note: Model 1 adds the contextual variable as main effect. Model 2 adds a three-way interaction term between external school evaluation reforms * parents' education * the additional contextual variable. Pooled analyses for men and women. The seven degrees of freedom difference between the models result from the interaction term between the additional contextual variable and external school evaluation (1 df), the interaction terms between three parental educational levels * the additional contextual variable (3 df), and the three-way interaction terms for three parental educational levels * external evaluation * the additional contextual variable (3 df).

Stronger effects in early tracking systems



Conclusions

- Somewhat surprisingly, we find that the introduction of external school evaluation policies, typically implemented through national Education Inspectorates, has negative effects on educational attainment for children of low SES. Little effects on medium and high SES groups.

Discussion

- We know nothing about schools in our data.
- While we touched upon three explanations, a better understanding would require more knowledge on school practices in the context of inspectorate visits.
- Current literature is not very analytical on this issue, but a few school-fixed effects studies have been done (mainly in England and Netherlands), with mixed results.

A photograph of a classroom with several children in the foreground, their hands raised in the air. The children are seen from behind, and their arms are extended upwards. The background is a dark chalkboard. The overall image has a semi-transparent dark overlay.

Thank you

<http://educsys.uva.nl>