The Effect of Treatment on the Untreated: Free Primary Education in Kenya

(Latest version can be found here)

William Labadie

December 5, 2017

PRELIMINARY - DO NOT CITE

1 Introduction

In this paper, I attempt to estimate the effect of expanding access to education on achievement among students directly unaffected by the expansion. Expanding access to education causes students previously unable or unwilling to enroll to come to school, potentially changing the composition of students, increasing the size of the student body and class pupil-teacher ratios. These changes may affect performance for students that would have enrolled in the absence of the policy change. In particular, I estimate the impacts of Kenya's Free Primary Education program on literacy rates for different ethnic groups. I use cohort variation in exposure to the policy to estimate the impact of the FEP program on student literacy, and I use an individual's ethnicity to proxy for his likelihood of being a student regardless of program exposure. I find that for individuals belonging ethnic groups that were relatively unaffected by the policy (that is, groups that appeared to have total access to primary education before the policy change), exposure to the policy has no significant bearing on likelihood of literacy. This preliminary finding suggests that exposure to the FEP does not significantly affect achievement for students that would have enrolled in the absence of the policy, or that the effects are not drastic enough to affect literacy rates.

The paper proceeds as follows. Section 2 describes the theoretical mechanisms through which the FEP could impact students that aren't directly affected; Section 3 describes the policy change and its context; Section 4 describes the data I use; Section 5 describes my first set of analyses and their results; Section 6 describes my second set of analyses and their results; and Section 7 discusses where to go from here.

2 Conceptual Framework

When a country expands access to education, it expands the set of students that enroll in a given affected grade. Certain groups of students that couldn't or wouldn't previously

attend the grade will enroll after the policy change, and this will affect the composition of students. For example, if a given grade becomes free to parents of students, students from lower-income families that were unlikely to have enrolled before the policy change will be more likely to enroll after the policy change due to the lowered cost of enrollment and the average family income of a student in that grade will decrease. Consistent with this hypothesis, Grogan (2009) and Nishimura, Yamano and Sasaoka (2008) find that Uganda's FPE program led to earlier and higher primary enrollment among low-income children. If peer effects exist, the performance of students that would have enrolled in the absence of the policy change will be affected by a changed peer group if the new students perform differently on average. Duflo, Dupas and Kremer (2011) exploit an experiment tracking students by ability to investigate the existence of peer effects among primary school students in Kenya, among other things; the authors find a positive direct effect of high-achieving peers on student achievement.

If infrastructure and staffing investments are not made by school administrators, average class sizes will increase with total enrollment after an access expansion. Class size is a popularly-discussed input of student achievement and teaching effectiveness. Hoxby (2000) exploits exogenous variation in primary school class sizes in the United States and finds no significant effect on student achievement. However, if increased class sizes are accompanied by a more heterogeneous classroom distribution of student abilities, instructors will be less able to target classroom instruction effectively, and student achievement will suffer (Duflo, Dupas and Kremer, 2011). Bosworth (2014) demonstrates that class size reductions affect heterogeneous North Carolinian fourth- and fifth-graders differentially. In particular, low-ability students benefit from class size reductions while high-ability students are relatively unaffected, and smaller classes tend to have smaller achievement gaps. Increasing pupil-teacher ratios are likely to hurt low-ability student performance.

The change in the composition of primary students brought about by the introduction of Kenya's FPE program seems likely to be detrimental to students that would have enrolled in primary school in the absence of the program. Students that enroll as a result of the program introduction seem likely to come from relatively low-income and low-education households, and likely to perform less well than the pre-FPE average. Lucas and Mbiti (2012) finds that public school students' parents are significantly less literate on average after the introduction of the FPE program. Further, high-ability or richer students may choose to enroll in private schools instead of public schools after the policy change (Hsieh and Urquiola, 2006). Peer effects, then, are likely to negatively affect achievement among students that would have enrolled in the absence of the policy. Lucas and Mbiti (2012) show that the FPE program significantly increased the number of students per teacher in Kenyan primary classrooms. This will negatively affect achievement among low-ability students that would have enrolled in the absence of the policy, but if new students lower the class average student ability, teachers may target instruction nearer to low-ability student levels. If high-ability students do flee to private schools, the distribution will further shift towards the low-ability students. That is, the increase in class sizes may not negatively affect the achievement of low-ability students that would have enrolled in the

absence of the policy change. However, the increase in class size seems likely to harm the achievement of middle-ability public primary students that would have enrolled in the absence of the policy change.

3 Background

The Kenyan government introduced its Free Primary Education (FPE) program in January 2003, eliminating all fees for public primary schools. Prior to the implementation of this program, public and private primary schools charged fees. On average, public school fees were approximately US$16 per year in 1997 (The World Bank, 2004), but some schools charged up to US$350 per year (Lucas and Mbiti, 2012). Fees were collected to pay for tuition, supplies, books, and facilities maintenance. After the introduction of the FPE policy, public primary schools received approximately US$14 per pupil per year from the government and parents of primary-age students faced no enrollment cost.

4 Data

I use data from the 2014 round of the Kenya Demographic and Health Surveys (DHS) to analyze the introduction of Kenya's FPE program. DHS is a nationally representative survey including individual-level data on an abundance of factors including educational attainment, ethnicity, age, domestic violence, and residence characteristics. For this project, I use the women's survey, which includes data on 31079 Kenyan women aged 15-49. I employ data on literacy to proxy for individual-level returns to education; educational attainment; ethnicity to proxy for intensity of treatment exposure; and year of birth to determine exposure to FPE.

Students born in 1990 or later were 13 or younger in 2003, the time of FPE introduction. Kenyan primary schools include grades 1-8, so children born in 1990 would have been able to take advantage of newly free 8th grade in 2003; children born in 1991 would be able to take advantage of free 7th grade in 2003 and free 8th grade in 2004; and so on. That is, the program's impact on students should be increasing over time. Figures 1, 2, and 3 confirm this prediction. Figure 1 presents mean birth cohort educational attainment for individuals born prior to 1996.[1] While average birth year educational attainment was rising over time prior to the policy introduction, there is an apparent discontinuity in educational attainment at birth year 1990, indicating that the FPE did expand access to education. Figure 2 shows the average attendance of at least some primary school by birth year for all individuals in my sample. Average primary school attendance is substantially higher for individuals born after 1990, and is increasing for each birth year cohort starting in 1991. Figure 3 shows average literacy rate by birth year. While each cohort's literacy rate is increasing over time prior to the introduction of the FPE policy, the rate of increase quickens for birth cohorts exposed to the FEP policy. This suggests

[1] Children born in 1996 would be 18 in 2014, the period of data collection. The educational attainment of children born after 1995, then, is right-censored in that none of them can report having completed secondary school or higher.

that the expanded access to education may have had an effect on literacy rates.

Kenya's FPE program, in expanding access to education, changed the composition of Kenya's student population. Table 1 presents the ethnic makeup of individuals of different levels of educational attainment before and after the introduction of FPE. The Kikuya group, which dominated the primary and secondary school attainers before the policy, saw their share of each attainment population fall significantly after the FPE introduction. The Mijikendu/Swahili and Somali ethnic groups, on the other hand, saw their shares of the primary and secondary attainer populations increase significantly after the policy change. It seems likely that, rather than reducing the number of Kikuya students enrolled in primary and secondary schooling, the policy increased the number of Mijikendu/Swahili and Somali students, increasing their shares of the student population and decreasing the share of Kikuya students.

5 Regression kink

5.1 Empirical specification

For a first look at the impact of Kenya's FPE program on literacy rates, I use a regression kink design (RKD). I estimate the change in slope at the birth year of first exposure to FPE, 1990. In particular, I am interested in the following structural model.

$$I(Literate)_{ic} = \alpha + \beta I(Primaryorhigher)_i + \rho(YOB - 1990)_c + u_{ic} \tag{1}$$

In this model, $I(Literate)_i$ equals 1 if individual i belonging to birth cohort c is able to read a full sentence and equals 0 otherwise; $I(Primaryorhigher)$ equals 1 if individual i attended at least some primary school and equals 0 otherwise; and $(YOB-1990)_i$ equals the distance from individual i's birth year to 1990, the year of the kink.

The first stage of this model is

$$I(Primaryorhigher)_{ic} = \pi_1 + \pi_2(I(YOB \geq 1990) \times (YOB - 1990))_c + \pi_3(YOB - 1990)_c + e_{ic} \tag{2}$$

and the reduced form of the model is

$$I(Literate)_{ic} = \tau_1 + \tau_2(I(YOB \geq 1990) \times (YOB - 1990))_c + \tau_3(YOB - 1990)_c + v_{ic} \tag{3}$$

. The regression kink design exploits an exogenous change in slope in the relationship between assignment to treatment a forcing variable x_f and a treatment variable D to estimate the causal impact of D on a variable on interest y. In this case, the introduction of FPE causes a change in exposure to primary school (D) at birth year (x_f)1990. If literacy (y) isn't independently affected by birth year, or has a smooth relationship with birth year, the regression kink framework will allow identification of the effect of primary school exposure on literacy.

I am interested in the effect of primary schooling on literacy for students that weren't directly affected by the introduction of FPE - students that would have enrolled in the absence of a policy change. To proxy for student likelihood of enrollment, I use ethnicity. In particular, I estimate the above model for Kikuya and Somali individuals seperately as well as for the full sample. The Kikuya were relatively likely to enroll prior to the policy and thus I assume they were relatively unaffected by the introduction of FPE; the Somali were relatively unlikely to enroll prior to the policy and the share of primary-educated Somali individuals increased for birth cohorts exposed to the FPE, so I assume they were relatively affected by the introduction of FPE. If Kikuya primary school enrollment is unaffected by the policy change, there will be no kink in primary school attainment at birth year 1990. If Kikuya literacy is affected by the changed composition of students, there will be a kink in literacy rate at birth year 1990. In this case, if the effect on literacy is negative, the RKD estimator β will yield a negative estimated impact of exposure to primary education after the introduction of FPE on Kikuya individuals.

5.2 Results

Figure 4 shows the first-stage regression kink for Kikuya women. Primary school attainment is high long before the FEP introduction, and rising over time. There is a modest discontinuity at the first birth cohort exposed to FPE, and a modest change in slope. This suggests that the policy change likely did not substantially affect Kikuya primary school enrollment behavior. Figure 5 shows the reduced-form regression kink for the Kikuya. There is no apparent discontinuity at 1990, and a very modest decrease in the slope. This suggests that the policy change likely did not substantially affect Kikuya literacy rates. Figures 6 and 7 show the first-stage and reduced-form results for the Somali ethnic group respectively. There is a clear discontinuity and increase in slope in both cases, suggesting that exposure to FPE substantially increased both primary school enrollment and literacy rates among Somali women at an increasing rate. Figures 8 and 9 plot the first-stage and reduced-form regression kinks for my full sample.

Table 2 presents the results of estimating equations (2) and (3) for Kikuya women, Somali women, and the full sample. The results confirm the relationships presented in Figures 4-9. Among the Kikuya, being born in or after 1990 does not significantly change the rate of primary attainment or the literacy rate. This suggests that exposure to the policy did not likely have a significant effect on Kikuya literacy rates. The policy may still have had indirect impacts on Kikuya returns to education - test scores, for example - but not on literacy. Being born in or after 1990 does have a positive and significant impact on the rate of primary attainment and on literacy rates among Somali women and on aggregate.

6 Student Likelihood Score

6.1 Empirical Specification

For this analysis, I attempt to better identify individuals that would have received primary education in the absence of the FPE introduction. To do so, I estimate the following

regression specification for individuals born prior to 1990 using OLS:

$$I(Primaryorhigher)_e = \xi + \gamma_e + \epsilon_e. \quad (4)$$

This regression will yield a "student likelihood score" γ for each ethnicity (indexed by e).

I then use these scores to estimate the relationship between primary school enrollment after the introduction of FPE, student likelihood score, and literacy using the following regression specification:

$$I(Literate)_{ice} = \sigma_1 + \sigma_2(I(YOB \geq 1990)) \times StudentLikelihoodScore)_{ce} + \sigma_3 I(YOB \geq 1990)_c + \sigma_4 StudentLikelihoodScore_e + \upsilon_{ice} \quad (5)$$

If the changed composition of students resulting from FPE affects the literacy rates of pre-FPE likely students, the exposure effect may be different for individuals of different birth years. The composition of students may not be very different in 1990 compared to 1989, for example. To address this, I estimate different exposure effects for each birth year:

$$I(Literate)_{ice} = \mu + \sum_{x=1990}^{1998} \delta_x (I(YOB) \times StudentLikelihoodScore)_{ce} + \sum_{y=1990}^{1998} \lambda_y I(YOB)_c + \kappa StudentLikelihoodScore_e + \omega_{ice}. \quad (6)$$

6.2 Results

I report the results of estimating equation (5) in Table 3. The Student Likelihood Score estimates suggest that exposure to the FPE policy does have a negative impact on students that were likely to attend school before the introduction of FPE. For a Kikuya woman (Student Likelihood Score of 0.989) born after 1990, for example, policy exposure reduces her likelihood of literacy by 15.5 percent. The linear probability model roughly suggests that this woman is 90.5 percent likely to be literate. For a Somali woman (Student Likelihood Score of 0.123) born after 1990, policy exposure reduces her likelihood of literacy by 1.9 percent. the model predicts that this woman is 27.1 percent likely to be literate.

The results of estimating equation (6) are displayed in Figure 10. The results suggest that the effect of exposure to FPE on literacy is increasingly negative for individuals that were likely to be students before the policy change. However, the coefficients on the year-of-birth dummies follow a mirrored path. The model predicts that a Kikuya woman born in 1998 is 95.1 percent likely to be literate, and that a Somali woman born in the same year is 10 percent likely to be literate. The true means are 0.989 and 0.473, respectively.

7 Some issues and next steps

The results of the regression kink estimation performed above suggest that exposure to Kenya's FPE had no significant effect on the literacy rates or the rate of primary school enrollment for students unlikely to be directly affected by FPE. However, the results of the linear probability model employing Student Likelihood Scores suggest that policy exposure had a negative effect on students that were unlikely to be directly affected by the policy. The differening results are probably a result of mis-identification.

A critical weakness of this paper is the use of ethnicity alone to measure an individual's likelihood of being a student before the introduction of FPE. DHS contains little demographic data on individuals that would be applicable to an individual's likelihood to enroll in primary school - things like childhood region of residence, parents' education level, parents' income, etc. Without further demographic controls, my analyses simply compare cohorts and ethnicities - there is actually no individual level to anything done in this paper. The Student Likelihood Score measure is certainly biased by omitted variables. The above "wishlist" of demographic characteristics are all likely correlated with an individual's ethnicity and primary school enrollment.

Another weakness of the paper is the use of literacy to proxy for student achievement. In populations that are likely to go to school prior to the introduction of FPE, even if the policy has an indirect effect on the quality of their schooling, the effect is unlikely to extend to their literacy. A measure such as test scores would be more appropriate, and more likely to reflect the true impact of the policy on those relatively unaffected by it.

I should also finish the RKD work: estimating the structural equation (1) using birth year to instrument for attendance of primary school. This RKD estimate β will only be unbiased if literacy has a smooth relationship with year of birth. The instrument's exclusion restriction does not strike me as particularly likely to be valid: one's year of birth likely affects his literacy through channels other than his likelihood of primary school enrollment. I should also investigate confounding changes that would differentially affect primary school enrollment and literacy rates of cohorts born after 1989.

Table 1: Ethnic Composition by Educational Attainment and Exposure to Policy

	Highest level of education attained								
	None			Primary			Secondary or higher		
	Pre	Post	Diff.	Pre	Post	Diff.	Pre	Post	Diff.
Embu	0.000	0.000	-0.000	0.016	0.007	-0.008	0.019	0.016	-0.003
			(0.000)			(0.002)***			(0.003)
Kalenjin	0.049	0.037	-0.013	0.158	0.168	0.010	0.148	0.159	0.011
			(0.008)			(0.008)			(0.008)
Kamba	0.014	0.006	-0.008	0.121	0.096	-0.025	0.099	0.105	0.006
			(0.004)**			(0.007)***			(0.007)
Kikuya	0.012	0.004	-0.007	0.184	0.095	-0.089	0.261	0.196	-0.064
			(0.003)**			(0.007)***			(0.009)***
Kisii	0.005	0.003	-0.002	0.052	0.044	-0.008	0.098	0.076	-0.006
			(0.002)			(0.005)*			(0.006)***
Luhya	0.024	0.007	-0.017	0.127	0.147	0.021	0.128	0.138	0.010
			(0.004)***			(0.008)***			(0.007)
Luo	0.010	0.002	-0.009	0.117	0.128	0.011	0.094	0.114	0.020
			(0.002)***			(0.007)			(0.007)***
Maasai	0.063	0.107	0.044	0.013	0.016	0.003	0.009	0.012	0.004
			(0.013)***			(0.003)			(0.002)
Meru	0.011	0.013	0.002	0.066	0.047	-0.018	0.058	0.053	-0.006
			(0.005)			(0.005)***			(0.005)
Mijikendu/Swahili	0.109	0.060	-0.049	0.050	0.090	0.040	0.022	0.039	0.017
			(0.011)***			(0.006)***			(0.004)***
Somali	0.309	0.309	0.000	0.009	0.022	0.014	0.008	0.027	0.019
			(0.007)			(0.003)***			(0.003)***
Taita/Taveta	0.002	0.0000	-0.002	0.019	0.010	-0.010	0.019	0.013	-0.006
			(0.001)**			(0.002)***			(0.003)**
Turkana	0.100	0.010	-0.001	0.009	0.024	0.015	0.005	0.008	0.003
			(0.013)			(0.003)***			(0.002)*
Samburu	0.099	0.120	0.021	0.005	0.015	0.010	0.003	0.004	0.002
			(0.014)			(0.003)***			(0.001)
Pokomo	0.008	0.009	0.001	0.013	0.020	0.007	0.005	0.006	0.001
			(0.004)			(0.003)**			(0.002)
Iteso	0.005	0.000	-0.005	0.011	0.014	0.003	0.006	0.009	0.003
			(0.001)***			(0.003)			(0.002)
Boran	0.050	0.044	-0.006	0.007	0.014	0.007	0.004	0.005	0.001
			(0.009)			(0.002)***			(0.002)
Gabbra	0.030	0.045	0.015	0.001	0.002	0.001	0.001	0.002	0.001
			(0.009)*			(0.001)			(0.008)
Kuria	0.004	0.002	-0.003	0.008	0.011	0.002	0.001	0.005	0.003
			(0.002)			(0.002)			(0.001)**
Orma	0.016	0.018	0.002	0.001	0.003	0.001	0.000	0.000	0.000
			(0.006)			(0.001)			(0.000)
Mbere	0.011	0.034	0.022	0.001	0.002	0.001	0.001	0.001	0.000
			(0.007)***			(0.001)			(0.001)
Rendille	0.015	0.021	0.006	0.001	0.002	0.001	0.000	0.001	0.001
			(0.006)			(0.001)			(0.001)

Notes: Standard errors are reported in parentheses. * denotes significance at the 90% level; **, significance at the 95% level; and ***, significance at the 99% level. Estimates represent the share of population held by a given ethnicity at a given education level. For example, of the survey respondents born prior to 1990 with secondary-level or higher educational attainment in 2014, 1.9% were Embu; of the respondents born in or after 1990 with secondary-level or higher educational attainment, 1.6% were Embu.

Table 2: Regression kink estimates

	Kikuya		Somali		Full sample	
	First stage: Primary or higher	Reduced form: Literacy	First stage: Primary or higher	Reduced form: Literacy	First stage: Primary or higher	Reduced form: Literacy
Slope change	-0.001	0.003	0.059	0.059	0.015	0.017
	(0.001)	(0.003)	(0.010)***	(0.009)***	(0.002)***	(0.002)***
N	4,463		1,491		26,342	

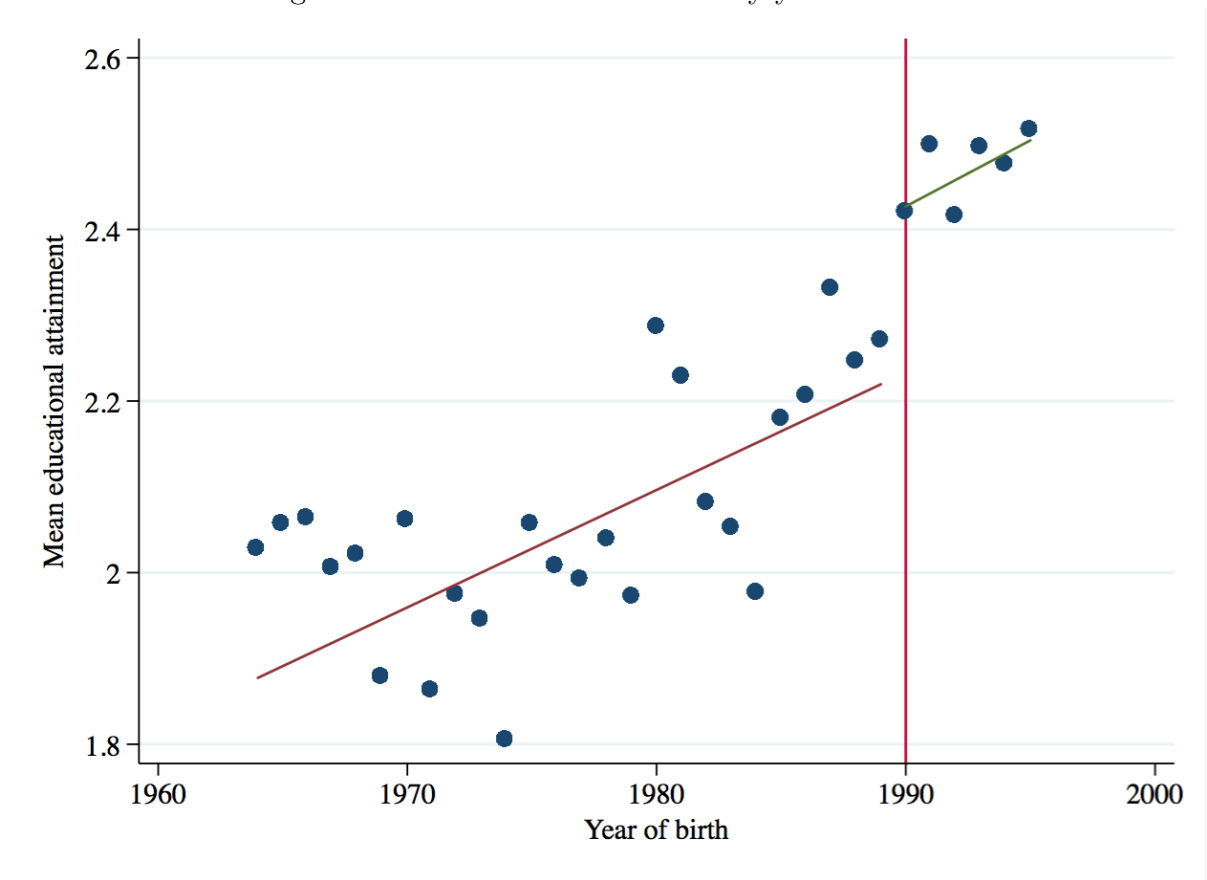
Notes: Standard errors are reported in parentheses. * denotes significance at the 90% level; **, significance at the 95% level; and ***, significance at the 99% level.

Table 3: OLS intensity-of-treatment estimates

	$I(Literate)$
$I(YOB \geq 1990) \times StudentLikelihoodScore$	-0.157
	(0.020)***
$I(YOB \geq 1990)$	0.241
	(0.018)***
$StudentLikelihoodScore$	0.889
	(0.010)***
Constant	-0.060
	(0.009)***

Notes: Standard errors are reported in parentheses. * denotes significance at the 90% level; **, significance at the 95% level; and ***, significance at the 99% level.

Figure 1: Educational attainment by year of birth



Notes: Attainment is a truncated variable that equals 0 if an individual received no formal education; 1 if he received some primary education; 2 if he completed primary school; 3 if he received some secondary education; 4 if he completed secondary school; and 5 if he was educated beyond the secondary level.

Figure 2: Mean primary attendance status by year of birth

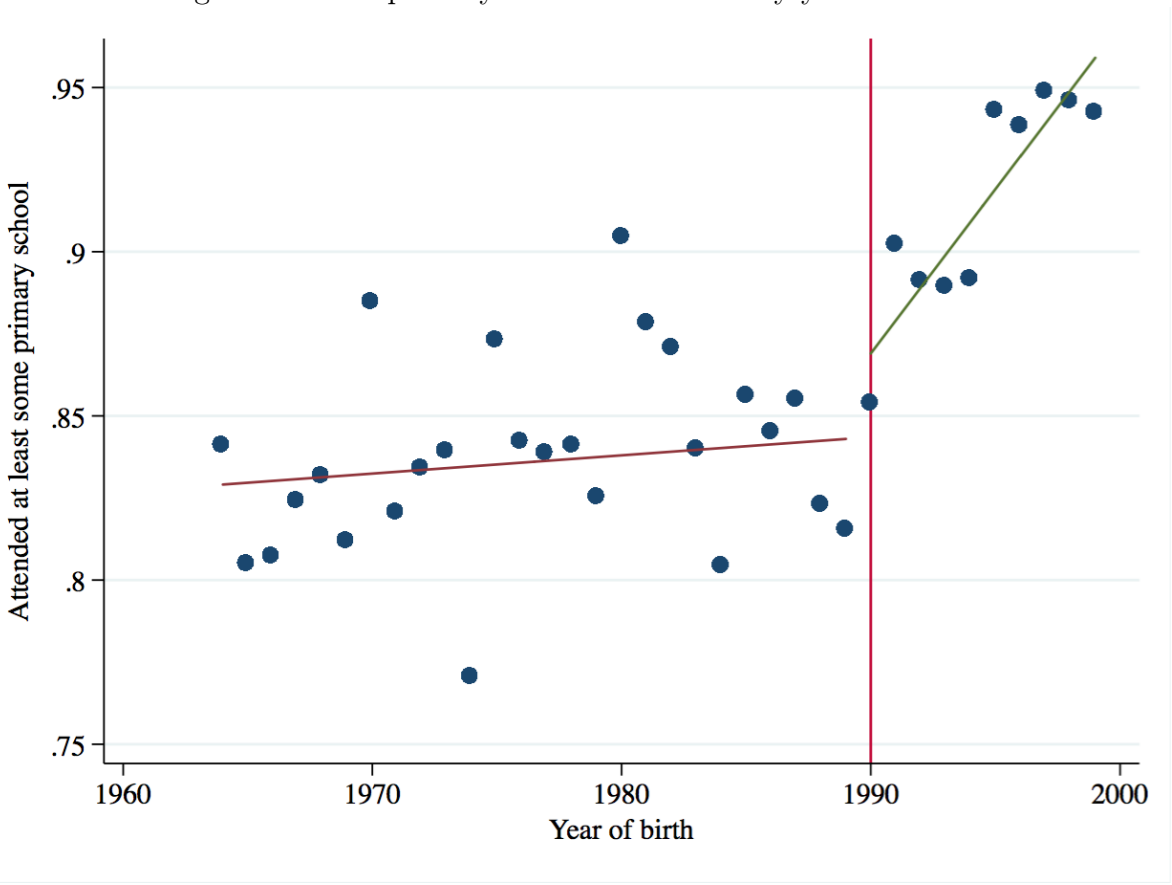


Figure 3: Mean literacy rate by year of birth

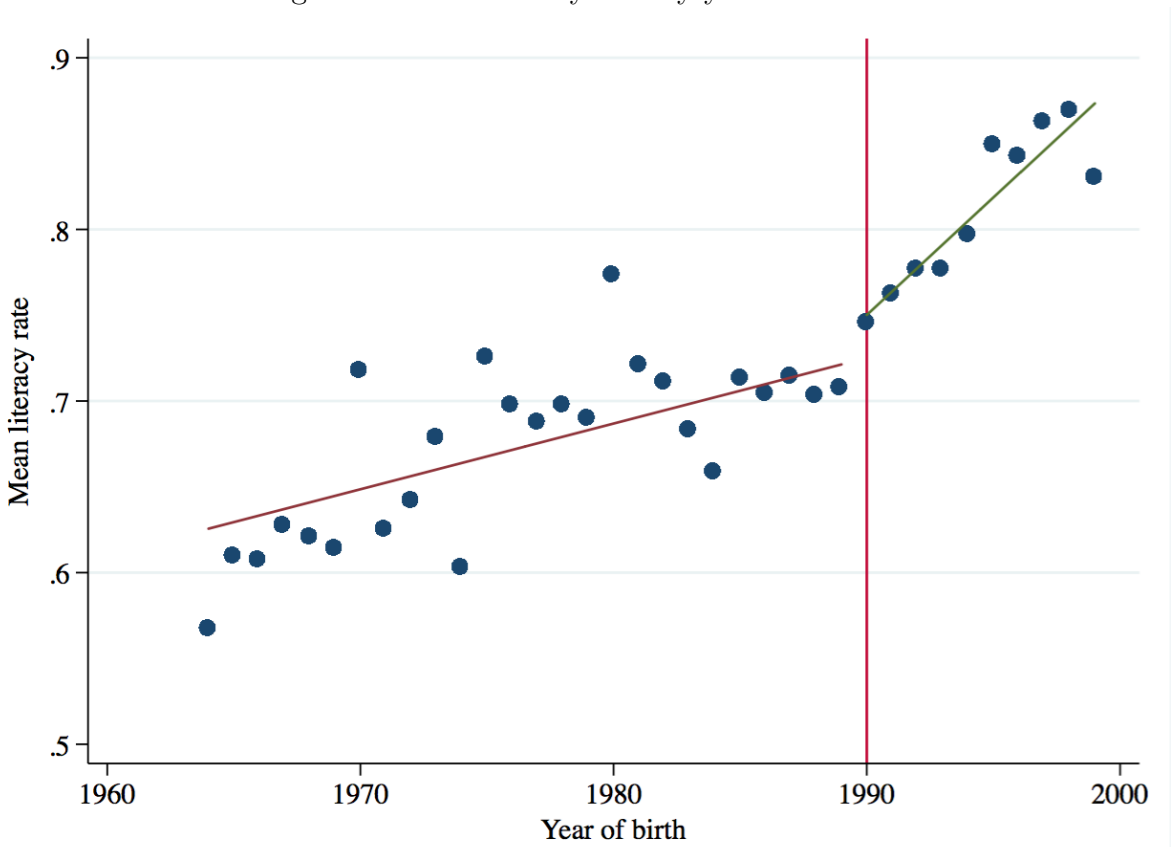


Figure 4: First stage: Kikuya

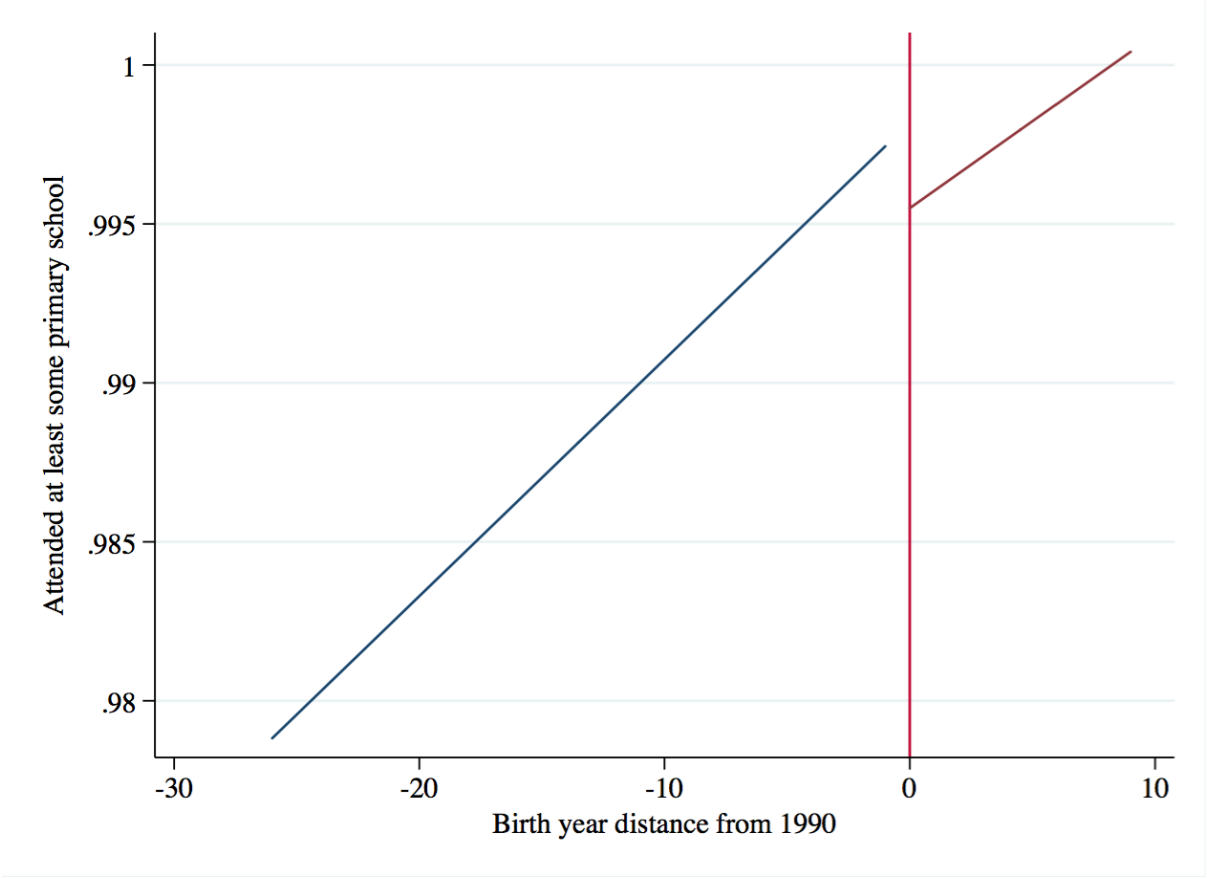


Figure 5: Reduced form: Kikuya

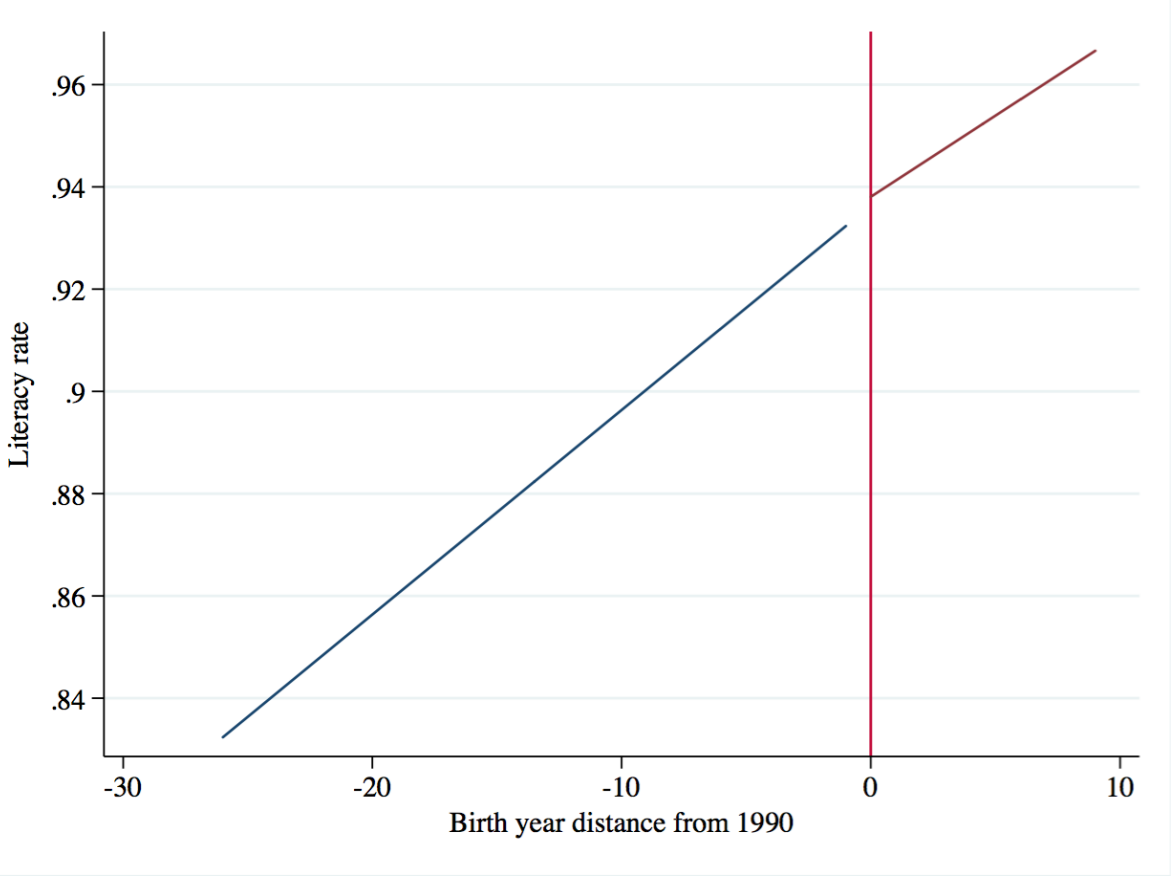


Figure 6: First stage: Somali

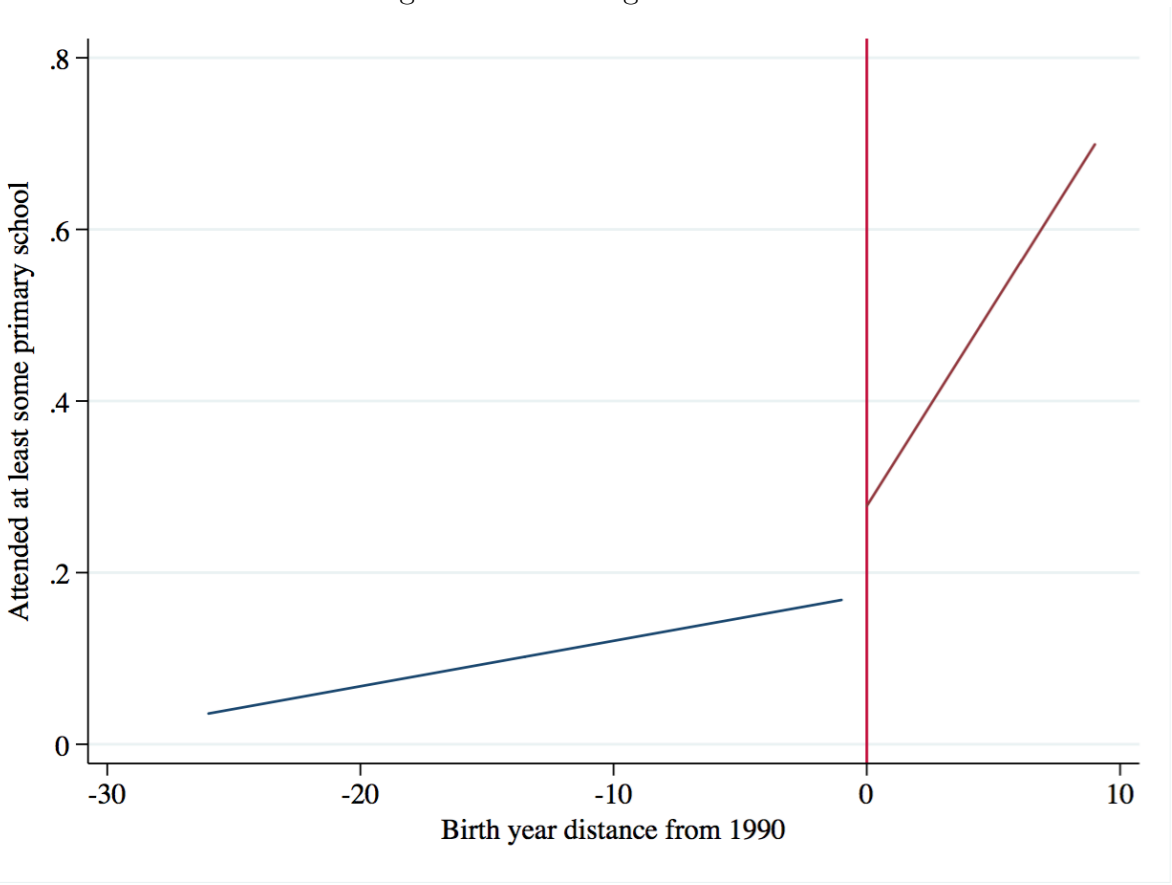


Figure 7: Reduced form: Somali

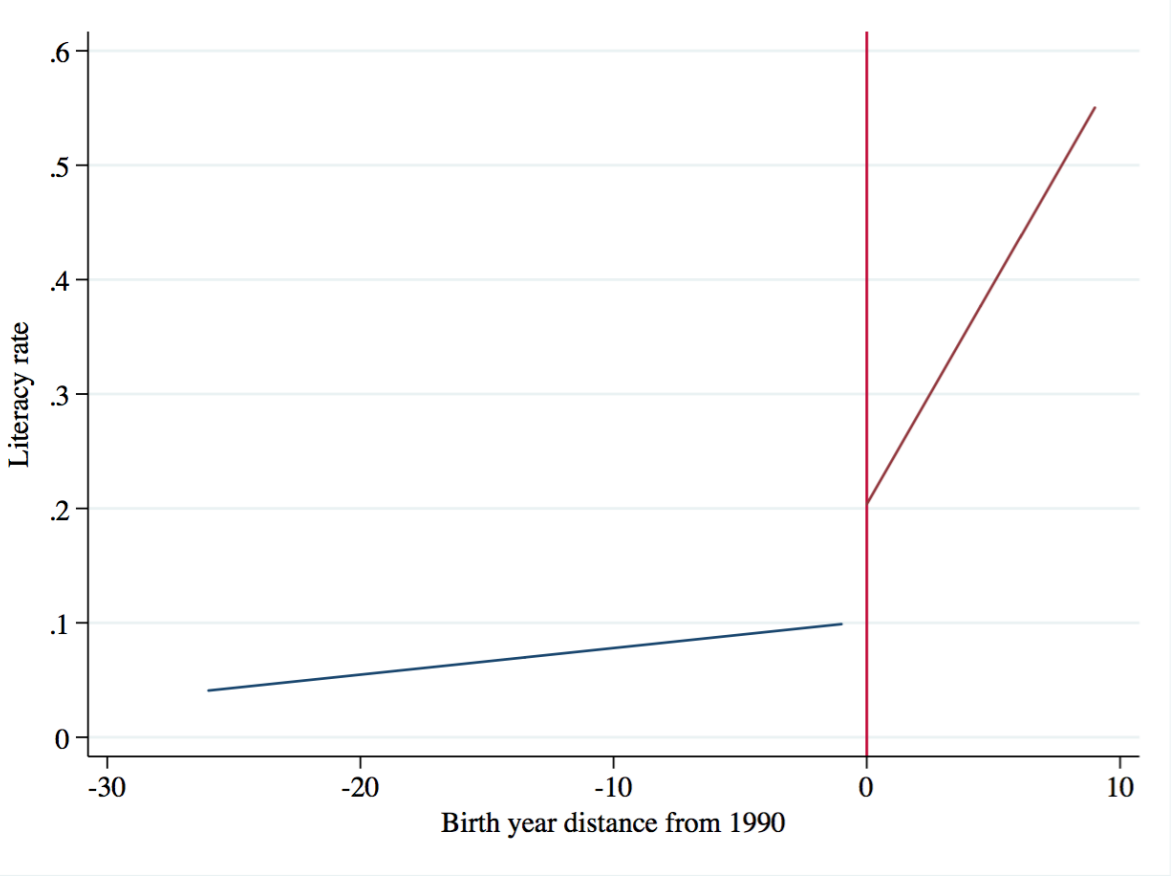


Figure 8: First stage: full sample

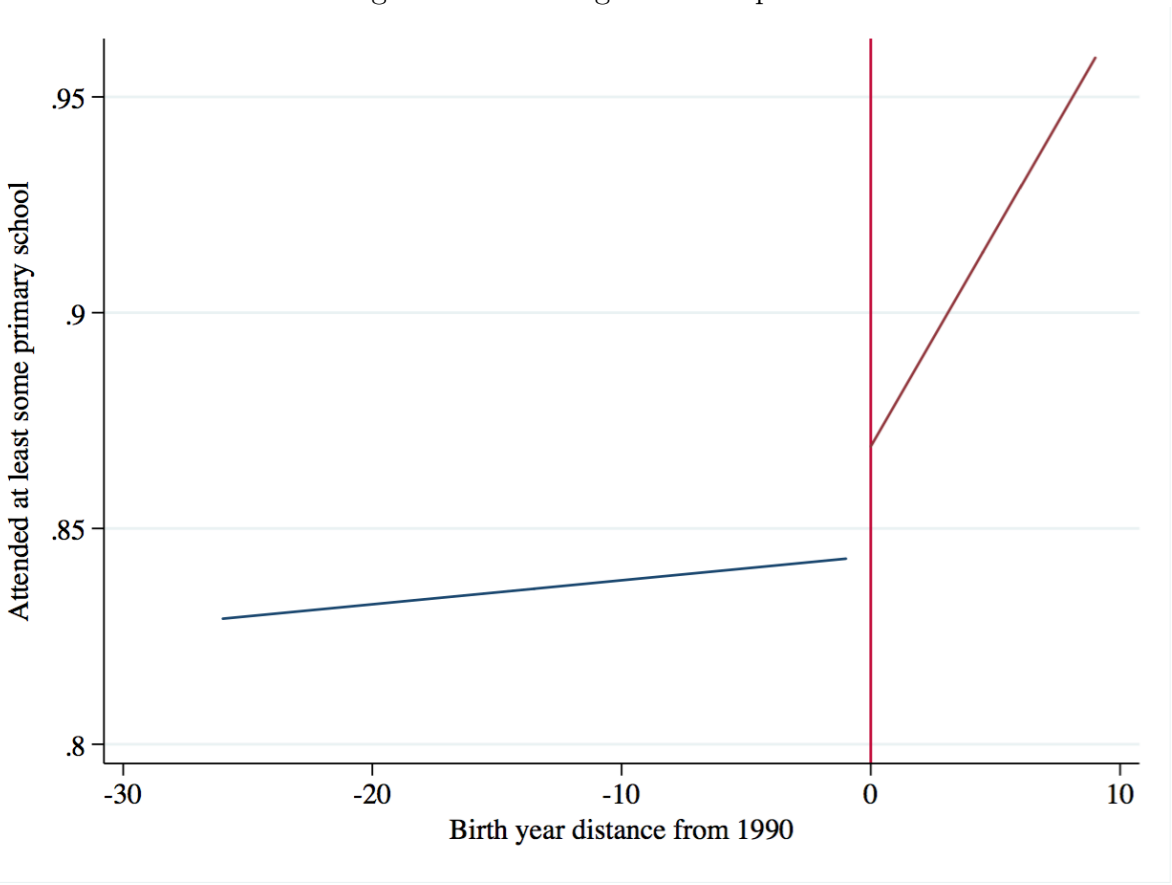


Figure 9: Reduced form: full sample

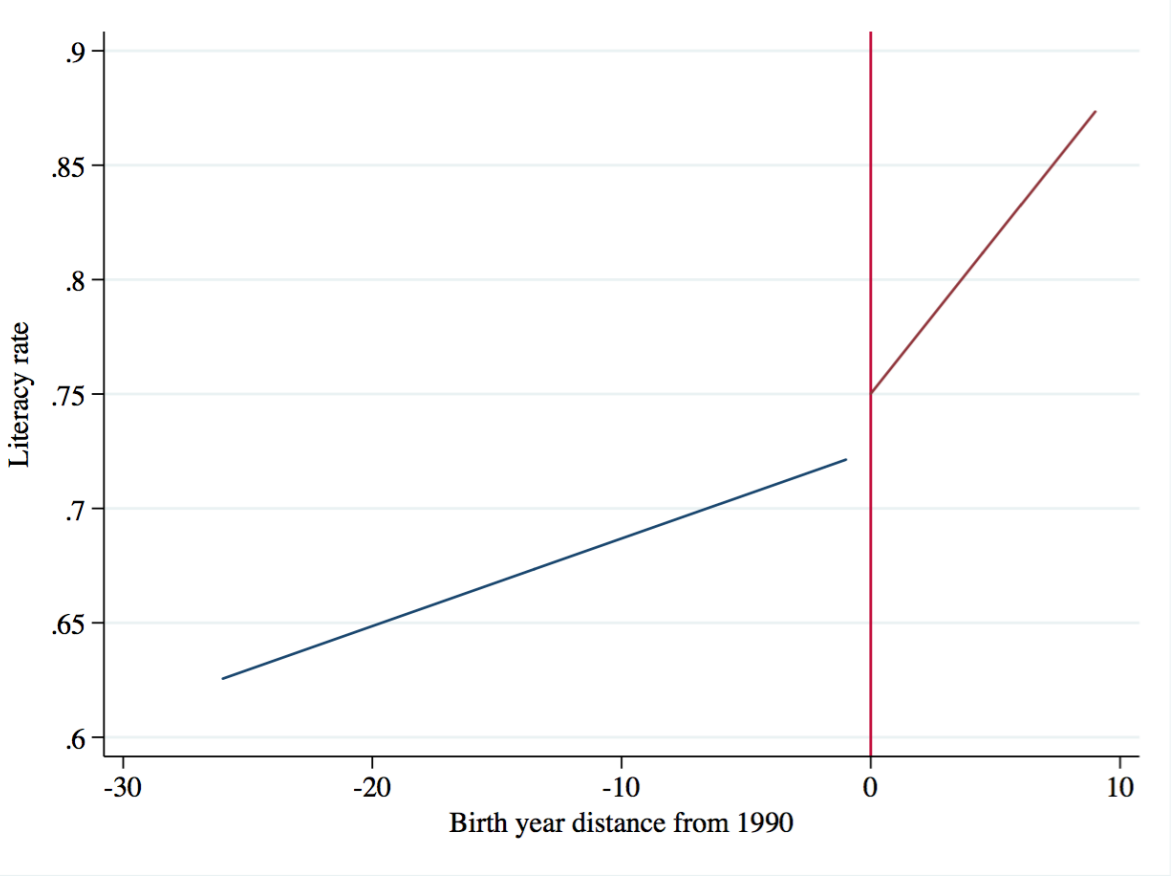
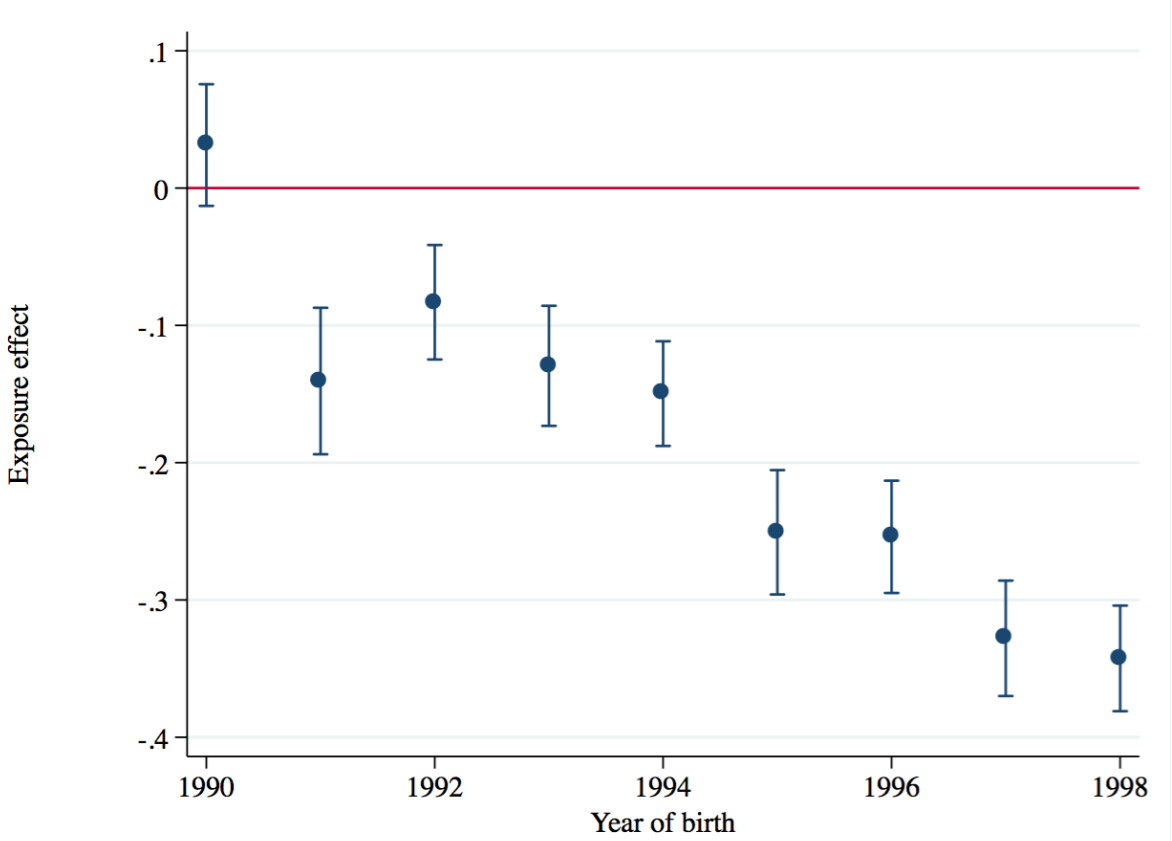


Figure 10: OLS coefficient estimates on $(I(YOB) \times StudentLikelihoodScore)$



Notes: Confidence intervals are at the 95% level.

References

Bosworth, Ryan. 2014. "Class size, class composition, and the distribution of student achievement." *Education Economics*, 22(2): 141–165.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Early Tracking: Evidence from a randomized Evaluation in Kenya." *American Economic Review*, 101(5): 1739–1774.

Grogan, Louise. 2009. "Universal primary education and school entry in Uganda." *Journal of African Economies*, 18(2): 183–211.

Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement : New Evidence from Population Variation." *The Quarterly Journal of Economics*, 115(4): 1239–1285.

Hsieh, Chang T., and Miguel Urquiola. 2006. "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program." *Journal of Public Economics*, 90(8-9): 1477–1503.

Lucas, Adrienne M, and Isaac M Mbiti. 2012. "The short-run effects of free primary education in Kenya." *American Economic Journal: Applied Economics*, 4(4): 226–253.

Nishimura, Mikiko, Takashi Yamano, and Yuichi Sasaoka. 2008. "Impacts of the universal primary education policy on educational attainment and private costs in rural Uganda." *International Journal of Educational Development*, 28: 161–175.

The World Bank. 2004. "Kenya: Strengthening the Foundation of Education and Training in Kenya - Opportunities and Challenges in Primary and General Secondary Education." The World Bank.