DOES THE LANGUAGE OF INSTRUCTION IN PRIMARY SCHOOL AFFECT LATER LABOUR MARKET OUTCOMES? EVIDENCE FROM SOUTH AFRICA

Katherine Eriksson[1]

ABSTRACT

This paper uses a change in the language of instruction in South African schools in 1955 to examine the effect of mother-tongue instead of English or Afrikaans instruction on long-term educational and economic outcomes. Using the 1980 South African census, a difference-in-difference framework allows me to estimate the effect of increasing mother-tongue instruction for black students from four to six years. I find positive effects on wages which I interpret as evidence of increases in human capital; these effects might have been larger in the absence of labour market discrimination against blacks under apartheid. I find positive effects on the ability to read and write, on educational attainment, and on the ability to speak English in predominantly English areas. I examine heterogeneous effects by region. This paper informs knowledge about the long-term effects of one aspect of a major apartheid education policy, the Bantu Education Act.

Keywords: South Africa, apartheid, Bantu Education Act, long-run effects

JEL Codes: I25, I28, N37

INTRODUCTION

The language of instruction in primary school and beyond is a subject of debate in post-apartheid South Africa where schools and parents must choose between the mother-tongue and English or Afrikaans for Black students, and in the United States in the context of Hispanic students. The language of instruction could directly influence language skill and literacy, but could also affect other important

1 Department of Economics, Orfalea College of Business, California Polytechnic State University. Email: keriksso@calpoly.edu.

Economic History of Developing Regions Vol. 29 (2) 2014
ISSN Print 2078-0389, Online 2078-0397

http://dx.doi.org/10.1080/20780389.2014.955272

outcomes, such as wages and educational attainment. To the extent that the language of instruction affects school quality it could also influence health outcomes, fertility, and social factors such as crime.[2]

Instruction in a student's native language could have both positive and negative effects. Educating students in their mother tongue for the first years of school will raise the chances of building non-language cognitive skills such as literacy and numeracy. In addition, school might be more accessible for students. Students could be more likely to enter school because they can understand the language. If they were going to drop out when the language of instruction changes to English, then increasing years of mother tongue instruction will prevent dropout.

There could also be positive school quality effects. For example, in South Africa, most teachers of Black children are Black so by teaching in English or Afrikaans they are teaching in their second language. Quality of instruction might increase if they teach in their first language. There is evidence that teachers who speak the same language as their students are also better able to impose discipline, increasing the probability that students learn well (Taylor & Vinjevold 1999). Finally, parents who speak only the mother tongue are more likely to be able to help with school work if their student's work is in a language they know.

On the other hand, teaching in English could increase English language skills, the language of most market transactions. In contemporary South Africa, using one language of instruction could enhance diversity in classrooms rather than separating students across language and therefore race. English skills are important for obtaining higher paid jobs, particularly in postcolonial countries where government jobs require the colonial language. There is a wealth of literature about returns to English speaking ability which is primarily based on the United States, but there is some evidence that there are high returns to English speaking abilities in South Africa as well.[3] Levinsohn (2007) calculates a return to speaking English for blacks between 40% and 80% in 1993 and 2000. Similar analysis of the 1980 census reveals only a 25% return as a result of labour market restrictions for blacks during apartheid.

Estimating the effect of the language of instruction is difficult because it is often an endogenous choice that parents can make on behalf of their children. In contemporary South Africa, parents can choose the school that their child attends and can also influence the language of instruction through voting for School Governing Bodies. Richer parents are more likely to send their children to schools with instruction in English. Variation in language of instruction in South Africa today is mainly in the first three years of school; mostly, former black schools use mother tongue instruction for this time period and former white schools almost

2 In this paper I look at literacy, educational attainment, and wages. The South African census does ask about disabilities but there are very few who report any disability in the 1980 census.

3 For the US, see Bleakley and Chin (2007), McManus et al. (1983), Chiswick (1991), and Dustmann (1994). In South Africa, Casale and Posel (2011) show a wage premium of around 40% for English reading and writing ability among blacks.

definitely use English or Afrikaans for the first three years. Nonetheless, even former black schools have increasingly started using English as the primary language of instruction from grade 1. Comparing outcomes of students who attend schools with mother tongue instruction versus English instruction would be extremely misleading as these two types of school are of vastly different quality.

This paper uses a difference-in-difference framework in the context of an exogenous change in the language of instruction to identify the effect of increasing a student's potential years of mother tongue instruction from four years to six years. In 1953, the new apartheid government in South Africa passed the Bantu Education Act which mandated eight years of mother tongue instruction for all Black students. Prior to the law, instruction was in the student's mother tongue for less than eight years. Before the policy change, mother tongue instruction lasted for four years in the Transvaal, Cape, and Free State provinces and for six years in the Natal province. I use the 1980 South African census to examine labour market and other outcomes of men and women aged 28 to 48 who were affected by the reform.

My results indicate that increasing exposure to mother tongue instruction had positive effects on male wages of 1.5–4%. I find that literacy and education levels increase in response to the policy change and argue that the wage increase is through these channels. In addition, I show evidence that English speaking ability increased as a result of the policy but only in predominantly English parts of the country. I also examine results for women and find larger effects on wages and educational attainment, consistent with the idea that mother tongue education increased accessibility more for girls than boys.

This paper is most closely related to Angrist and Lavy (1997) and Angrist et al. (2008). Angrist and Lavy (1997) use a natural experiment in Morocco to infer the effect of French language skills on test scores and the returns to school. Using a change from French to Arabic instruction after grade six, their difference-in-difference specification uses cohorts affected and unaffected by the change as well as students who obtained more than grade 6 and those who did not. They find strong negative effects of 'Arabization' on both test scores and returns to school.[4]

Angrist et al. (2008) use a similar identification strategy to determine the effects of Spanish-only schooling on language skills in Puerto Rico. They use the change in 1949 from some years of instruction in English to all instruction in Spanish. Controlling for education-cohort trends, they find that the language of instruction does not matter for English skills. Their paper highlights the importance of accounting for selection of skill groups into higher grades as well as controlling for cohort-schooling trends which can bias results.

4 An important difference between my paper and Angrist and Lavy's is that I focus on instruction in mother tongue. In contrast, a majority of the population of Morocco is Berber and does not speak Arabic or French as a first language so this policy probably did not affect more than the urban minority

Within the African context, there are multiple studies which have examined English versus native-language instruction. In South Africa, a recent paper by Taylor and Coetzee (2013) uses a school fixed effect strategy to show that students perform better on later grade English exams when taught in their mother tongue instead of English for the first three years. There is also some mixed evidence from Nigeria (Akinnaso 1993), Mozambique (Benson 2000), and Uganda (Piper & Miksic 2011) although these papers either rely on very small samples or non-causal methods.

My paper makes two main contributions to the literature on language of instruction. First, I use cross-province variation and over-time variation which allows me to construct difference-in-difference estimates without using educational attainment as a source of variation. Educational attainment is a potentially problematic source of variation because it is likely endogenous to the policy reform. If there was selection into higher grades induced by the policy change and these were less skilled students, then the estimates of the effect will be biased. Second, by using the 1980 census, I look at long run effects of this policy change; rather than looking at grades in subsequent grades, I look at the long-run effect on wages and completed educational attainment as well as literacy and migration outcomes.[5]

The paper is organized as follows. Section 2 provides background on South Africa's population and history, and outlines the language reform in detail. Section 3 presents the difference-in-difference estimation strategy. Section 4 explores the results and Section 5 concludes.

BACKGROUND

In South Africa, the Black, or African, population made up approximately 80% of the population in 1980. Whites constituted 9% which can be broken into Afrikaans- (65%) and English-speaking (35%) groups. The Coloured population constituted an additional 9% of the population and was found predominantly in the Cape Province.[6] The Indian population made up the remaining 2%. Until 1994, the Black population was not eligible to vote or influence policy; the laws

5 A third contribution is that the policy change in South Africa was exogenous to the population it affected. It was imposed by the white government and was not in response to popular demands. The policy changes in Morocco and Puerto Rico were the result of national debates about the best language of instruction.

6 In South Africa, 'Coloured' means mixed race. Most Coloured people are of Afrikaans and Black descent and speak Afrikaans as a first language. This race was not subject to all of the *apartheid* restrictions but did have a specific education department and set of schools. Limited voting rights and representation were given to this group in various forms throughout the apartheid era.

which constituted the system of apartheid were gradually introduced through the 1950s after the general election success of the mainly Afrikaner National Party in 1948. During the 1950s, formal separation of the races was legally established.

Mariotti (2012) argues that the differences in educational opportunities between less-educated Afrikaans and more-educated English speakers and the resulting income differential influenced the direction of policy in South Africa throughout the twentieth century. The Afrikaans-speaking population, much larger than the English population, dominated South African politics post-1948 but the English population controlled the government for the first half of the twentieth century. The post-1948 apartheid government's goal was to separate the lives of the races, with the Black population providing a labour source for unskilled and semi-skilled jobs. To this end, major policy changes included the restriction of movement of the Black population, the creation of Black 'homelands', formal job reservations for whites, and a reform of the education system in the early 1950s.[7]

The apartheid government worried about competition for semi-skilled jobs between Blacks and white Afrikaans speakers, so formal job reservation was legalized through the prohibition of Black union membership. Job skill classifications were used to determine which races could work in these jobs. Mariotti (2012) shows that this system had partially broken down by 1980 due to increased skills of whites and increasing demand for semi-skilled labour. Blacks were more likely to be working in semi-skilled jobs in 1980 than a decade previously, but there were still substantial restrictions on the jobs they could obtain. To the extent that these restrictions meant that skill was not rewarded in the labour market, any positive effects of the language reform on literacy would not necessarily have translated into higher wages.

Before 1950, the education system was a mix of publicly-provided education and mission or other private schools; provinces enjoyed autonomy over education policy. In 1951, the new government established the Eiselen Commission whose task was to examine the Black education system and make suggestions for a

7 Formal homelands were created in areas with land unsuitable for major agricultural production. The apartheid government envisioned that the Black population would live solely in these areas, which would eventually become nominally independent, and that Black workers would only serve as migrant labor. The homelands gradually obtained 'independence', with the Transkei, Bophutatswana, and Venda homelands being independent by 1980; an additional seven were still considered a part of South Africa. The Ciskei became independent in 1981. The four independent homelands as of 1981 are sometimes referred to as the TVBC states. Despite the nominal independence of the homelands, they were never recognized by the international community. Unfortunately, in 1980, the 'independent' homelands of the Transkei and Venda were not enumerated in the main census. Each independent homeland had their own census but only the one for Bophutatswana is available from Statistics SA.

national education policy. In addition to a revision of the curriculum, the commission recommended that all Black students be taught in their mother tongue for the first eight years of school.

In 1953, the apartheid government passed the Bantu Education Act which established a separate department of education for Black education as well as national control over all education policy. The objective of the Bantu Education Act was to standardize control of education across provinces. Some saw the new education system as a method of keeping blacks in a position of inferiority, but Giliomee (2009) argues that the Act placed a large emphasis on mass literacy to prepare youth for semi-skilled jobs. The education system was not meant to prepare youth for highly skilled or government jobs, except in the homelands where the government argued that the development of new governments would require a small skilled Black population. Prior to the restructuring of the education system, enrolment rates were low for Black students. In 1936, enrolment rates of seven- to 16-year-olds hovered between 15% and 30% (Hartshorne 1953). The Eiselen Commission's report set a goal of almost doubling primary school enrolment by 1959. In fact, enrolment at this level more than double by 1963 (Giliomee 2009).

The Act also implemented the Eiselen Commission's recommendation that all students be taught in their mother tongue for the first eight years of school (Taylor & Vinjevold 1999). Students were then taught after year eight in a mix of English and Afrikaans.[8] The predominant white language in the area was required as a subject from the second year of school and the other 'white' language was introduced in the fourth year (Hartshorne 1992).

Language of instruction policy before the 1950s was largely a relic of the practices of mission schools, which were established in the nineteenth century. Natal schools taught in the student's mother tongue for six years and in the three other provinces – the Transvaal, Cape, and Free State – instruction was in the student's mother tongue for only four years. All students were taught in either English or Afrikaans thereafter.

After the reform, the new language policy was implemented from the fifth year of school (grade five) to the eighth year of school (grade eight) gradually, beginning in 1955 (Taylor & Vinjevold 1999). In 1955, all fifth graders were taught in their mother tongue. In 1956, mother tongue instruction was extended to grade six. By 1958, all students through grade 8 were taught in their mother tongue. Outside of Natal the implementation of the policy meant that students in grade 5 or below in 1955 would have had eight years of mother tongue instruction but

8 Instruction was meant to be given in an equal proportion of English and Afrikaans after grade eight. In practice, this was not strictly enforced. The policy was enforced following the Afrikaans Medium Decree in 1974 which led to student protests which culminated in the Soweto Riot in 1976.

anyone who had passed grade 5 by 1955 would have had only four years. For example, a student who was in sixth grade or beyond in 1955 was already taught in English and so did not switch back to her mother tongue. In Natal this meant that the policy did not take effect until 1957 when grade 7 switched from instruction in English to mother tongue. Here the first cohort affected was those who were in fifth grade in 1955 and consequently in seventh grade in 1957. Therefore, anyone who had passed fifth grade by 1955 was not affected by the change in either the treatment or control provinces.

Hartshorne (1992) argues that the policy was strictly enforced. Other evidence is more mixed; Lafon (2009) documents interviews with individuals who remember being taught in their mother tongue after grade 8 as well as those who attended a school in which they switched to English before grade 9 during this time period. However, the national exam was given in native languages for the first time in 1959, so the policy was applied at least at a national level. In addition, Soudien (2002) documents interviews with teachers who claimed that part of their disenchantment with Bantu Education was due to increased inspections of schools which suggests that the new school superintendents did visit the schools and check on the policy's enforcement. I discuss possible threats to identification in the next section which could have to do with non-compliance with the law. I also discuss the implications of it possibly being enforced at a different time than the law stated.

Language policy was obviously not the only aspect of education policy affected by the Bantu Education Act. It is possible that school quality could have increased across the board as a result of more inspections of schools. Other evidence suggests that quality decreased due to the large increase in numbers of students enrolling in school without a matching increase in qualified teachers. Many other changes to the structure and curriculum also happened with this Act. My identification strategy outlined in the next section will depend on the assumption that any other changes during this time period did not affect the same grades at school (grades 4 to 6) differentially across treatment and control provinces.

Bantu Education was generally unpopular. Teachers disagreed with some of the policies and reacted in various ways, including community action and not coming to work (Soudien 2002). On average, parents opposed the mother-tongue policy, preferring their children to learn in English or Afrikaans as early as the beginning of school (Lafon 2009). The Black homelands which became independent in the 1970s and 1980s reverted immediately to four years of mother-tongue education (Giliomee 2009). Despite the unpopularity of the mother-tongue language policy, Kathleen Heugh (2000: 24) argues:

> Despite the cognitively impoverished curriculum, eight years of mother tongue instruction gave pupils time to learn their own language and to learn a second and a third language well sufficiently well to make the switch in medium in the ninth year. During the first phase of Bantu Education, 1953–1976, the matriculation results improved, despite the poor curriculum…

Table 1: Years of potential mother tongue instruction by cohort and birth province

	Born before 1942 ('Old')	Born in or after 1942 ('Young')	Difference:
TCF	4	8	8 – 4
Natal	6	8	8 – 6
		Difference in Difference:	6 – 4

ESTIMATION STRATEGY

My difference-in-difference strategy exploits different exposure to mother tongue instruction across cohorts and provinces. My approach is similar to that of Angrist and Lavy (1997) and Angrist et al. (2008) in that it uses a policy change as a natural experiment to estimate a difference-in-difference equation, but differs because it uses cross-province instead of cross-education group variation.

My difference-in-difference strategy is described in Table 1. I compare cohorts affected and unaffected by the policy change and also compare provinces which were affected in different ways. I define a variable TCF equal to one if an individual is born outside of Natal and zero otherwise; that is, it is equal to one if born in the Transvaal, Cape, or Free State provinces and zero if born in Natal. The 'old' cohort includes individuals born before 1942, while the 'young' cohort includes those born in or after 1942. The first difference is the difference across cohorts within each province. In TCF, the 'old' cohort was taught in their mother tongue for four years while the 'young' cohort had eight years of mother tongue instruction. Similarly, in Natal, the 'old' cohort received six years of mother tongue instruction, while the 'young' cohort experienced eight years. Therefore, the policy change induced an increase of four years of mother tongue instruction outside of Natal, in TCF, and only two years in Natal. The second difference is across provinces. Subtracting the change of two years from grade six to eight in Natal from the change of four years from grade four to eight in the Transvaal, my difference-in-difference estimator measures the effect of increasing mother tongue instruction by two additional years.

To assign individuals to groups, I must make certain assumptions. First, I assume that children started school at age eight and progressed through one year of school each year. The school starting age of eight is documented in government publications (South African Bureau of Racial Affairs 1955).[9] This assumption is strong in a developing country, so I present results relaxing this assumption so

9 Analysis of black school-aged children in the 1980 census shows that 16.5% of six-year-olds and 39% of seven-year-olds had already completed one year of school.

that students could be one to two years behind where I predict them to be in 1955 when the policy took effect. Second, I assume that children attend school in their province of birth which is reasonable given the low internal migration rates and migration restrictions in this time period. Third, I assume that the policy was implemented as written, starting in 1955 and being fully implemented by 1957. Finally, the policy must have been enforced; any non-compliance which was random with respect to the different provinces will serve to attenuate my estimates.

The validity of the difference-in-difference strategy hinges on the parallel trends assumption. This assumes that Natal and the other provinces had parallel pre-trends in the outcome variables of interest and that the trends would be the same in the absence of treatment. That is, cohort trends in the outcome variable would have been the same across treatment and control provinces in the absence of the policy change. Some evidence in favour of this assumption is provided in Figures 1 to 4. Figure 1 shows the average level of income by age for treatment and control groups. Figure 2 breaks this down by birth province. Figures 3 and 4 show the average level of education for the same groups. Treatment and control provinces track closely in the pre-period, but Natal has overall lower levels of education and income.

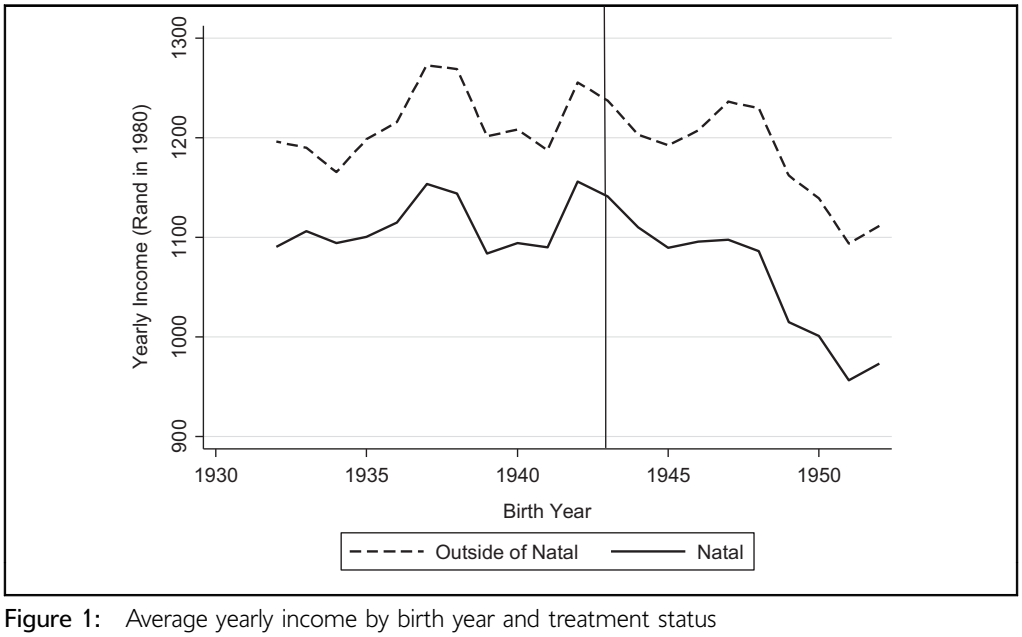

Figure 1: Average yearly income by birth year and treatment status
Notes: This figure plots a 3-year moving average of income by birth year and treatment status of birth province. Individuals born in Natal are compared to those born outside of Natal. The vertical line divides affected (right) and unaffected (left) cohorts. I restrict to men born outside of homelands which were not enumerated in the 1980 census. The 3-year moving average is used to smooth out noise in the data caused by misreporting of age in the census.

Figure 2: Average yearly income by birth year and birth province
Notes: This figure plots a 3-year moving average of income by birth year and birth province. The vertical line divides affected (right) and unaffected (left) cohorts. I restrict to men born outside of homelands which were not enumerated in the 1980 census. The 3-year moving average is used to smooth out noise in the data caused by misreporting of age in the census.

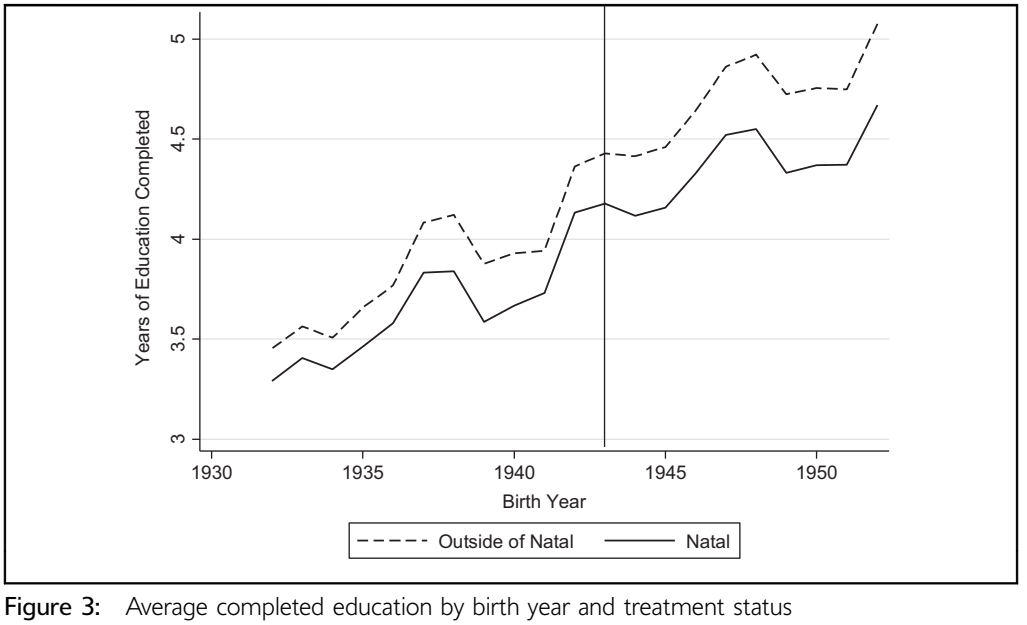

Figure 3: Average completed education by birth year and treatment status
Notes: This figure plots a 3-year moving average of education levels by birth year and treatment status of birth province. The vertical line divides affected (right) and unaffected (left) cohorts. I restrict to men born outside of homelands which were not enumerated in the 1980 census. The 3-year moving average is used to smooth out noise in the data caused by misreporting of age in the census.

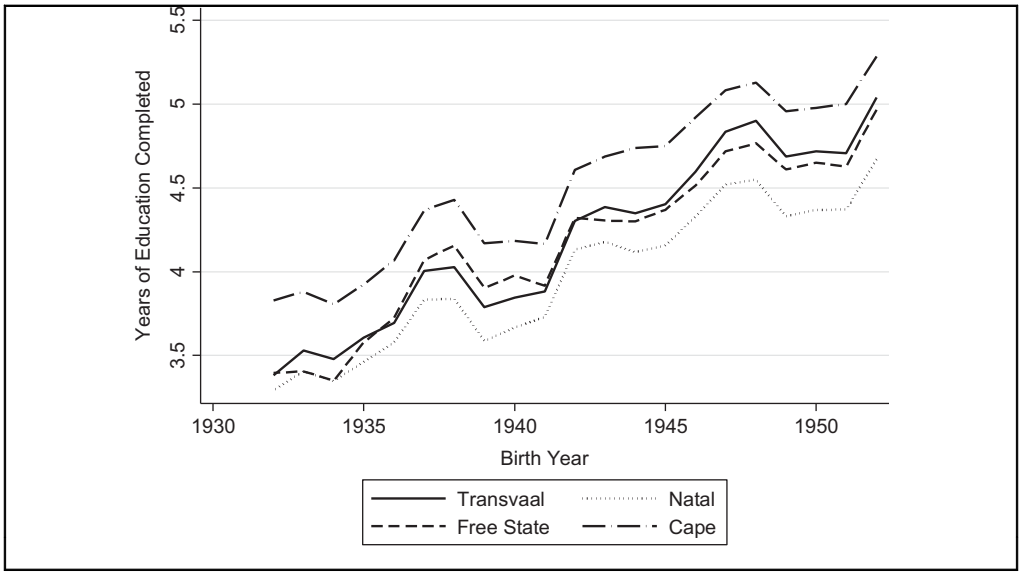

Figure 4: Average completed education by birth year and birth province
Notes: This figure plots a 3-year moving average of completed education by birth year and birth province. The vertical line divides affected (right) and unaffected (left) cohorts. I restrict to men born outside of homelands which were not enumerated in the 1980 census. The 3-year moving average is used to smooth out noise in the data caused by misreporting of age in the census.

My first specification compares individuals in the two cohorts of interest who were born in Natal to those born outside of Natal:

$$y_{ijm} = \alpha + \delta_1 young_j + \delta_2 TCF_m + \beta young_j * TCF_m + \theta_j + \gamma_m + \varepsilon_{ijm} \qquad (1)$$

for individual i born in year j in municipality m. The variables *TCF* and *young* are absorbed by fixed effects for birth municipality and for each age so are not reported in the results. The parameter β is the difference-in-difference estimator for the effect of increasing years of mother tongue instruction by two additional years from four years to six years. Standard errors are clustered at the level of birth province and p-values are calculated using a t-distribution with three degrees of freedom (Bertrand et al. 2004; Cameron et al. 2008; Cameron & Miller 2014).[10]

Note that these difference-in-difference estimates are intent to treat (ITT) effects, not average treatment effects (ATE) or treatment on treated effects (TOT). I do not require that any student actually attain grade 5 in order to be considered in the treatment group. However, if no one passed grade four, we would not expect to find economically significant results. From a policy perspective, we are most interested in the TOT effect; in the results section, I discuss a reasonable range for the TOT effect based on the proportion of students who reached high enough grades of school to be affected.

10 Standard errors are usually smaller if I cluster at the level of birth municipality.

Table 2: Percentage of men reporting each age by birth province

Age	Transvaal	Natal	Free State	Cape
28	7.84	7.87	8.03	7.85
29	6.16	6.25	6.09	6.36
30	9.03	8.89	8.48	8.56
31	4.76	4.75	4.62	4.69
32	6.54	6.3	6.37	6.36
33	4.38	3.76	4.19	4.65
34	4.66	4.38	4.28	4.68
35	5.74	5.64	5.82	5.4
36	5.15	5.34	5.00	4.79
37	3.37	3.33	3.37	3.77
38	5.66	5.75	5.47	5.37
39	4.11	4.33	3.93	4.67
40	7.21	7.33	7.26	7.64
41	2.88	2.53	2.77	3.09
42	4.12	4.4	4.08	4.14
43	2.72	2.68	2.66	2.8
44	2.59	2.5	2.34	2.6
45	4.6	5.07	4.83	4.16
46	2.6	2.64	2.97	2.64
47	2.2	2.3	3.71	2.36
48	3.67	3.97	3.73	3.44

CENSUS DATA AND DESCRIPTIVE STATISTICS

The analysis uses individual-level census data from a 100% universe of the 1980 South African census.[11] My dataset contains information about age, municipality

11 The full count of the 1980 census is available on Statistics South Africa's website. Additional censuses are available from 1970, 1985, 1991, 1996, 2001, and 2007 but 1980 is the only census which includes detailed information about place of birth as well as income for the Black sample. The 2001 census reports province of birth as defined in 2001, but the provincial boundaries changed after the end of apartheid so it is difficult to create a 1-1 mapping between these reports and where the individual would have been born and educated. The 2001 census also only reports income in intervals. The census asked about place of birth in 1970 but this is not available in the published data.

of birth,[12] province of residence, education level, yearly income, employment status, and the self-reported ability to speak, read or write in English, Afrikaans, and the individuals' home language. I restrict to black men and women between the ages of 28 and 48.

Multiple measurement problems in the Census could potentially affect the results. First, the data exhibit significant age heaping at multiples of five as well as at age 32. Individuals who report one of these ages have lower average educational attainment and lower reported income. This age heaping does not fall at age 38, the cut-off between young and old cohorts, so it will only affect the results if it is differential across provinces. Table 3 provides evidence that the age heaping is relatively consistent across provinces. Closer analysis of the age-heaping of income and education levels shows a similar pattern.

Second, 22% of blacks were not enumerated in the census (South African Data Archive 1980). If this under-enumeration was in areas with high returns to language skills (e.g., urban areas), then this will potentially affect my results if the extent of migration into these areas differed across provinces and cohorts. For example, if those who were more affected by the policy change were more likely to move to urban areas, then undercounting these individuals would lead to understating my results.

Finally, the Black homelands which had already attained 'independence' by 1980 were not enumerated, even though in 1955 they were affected by the policy change. Migrants out of these homelands were enumerated in the census, so if these migrants are different than non-migrants then the difference-in-difference coefficient could be biased. Therefore, I drop individuals born in the independent homelands (Transkei and Venda) which were not enumerated in the 1980 census and for which a separate census is not available. I also drop individuals born in Bophutatswana; even though this census is available separately in 1980, the variable definitions, particularly income, are not the same.[13]

A full list of the dependent and the independent variables used in the analysis are defined in Table 3. I restrict my attention to individuals who were between ages 28 to 48 in 1980 because younger cohorts were potentially affected by various disruptive incidents in schools in the 1970s, and older cohorts do not provide a good comparison group due to trends in educational attainment.[14] This age group is divided into young and old cohorts based on the age at the time of the language reform. As discussed above, the young cohort is defined as anyone who was born in or after 1942, or who is 38 or younger in 1980.

12 In 1980, there were 324 reported municipalities of birth. Including fixed effects for birth municipality allows me to control for any time-invariant differences in school quality at this level.

13 This shouldn't affect the results unless there is selective migration *into* these areas of people born outside of the areas which was induced by the policy.

14 Results are robust to changes in the age cut-off on either side (results not shown).

Table 3: Variable definitions

Independent Variables	
TCF	Born in Transvaal, Cape or Free State Provinces
Young	Born in or after 1942
Transvaal	Born in the Transvaal Province
Cape	Born in the Cape Province
Freestate	Born in the Free State Province
Dependent Variables	
Years of education	Number of years of education completed
Log Income	Natural log of reported yearly income
White Ability	Speaks or Speaks, Reads and Writes English or Afrikaans
Literate	Reads and Writes in any language
Migrant	Living outside of birth province in 1980
Employed	Reports being an Employer or Employee
Children	Number of children ever born (women)

Individuals are assigned to their province of birth based on the reported municipality of birth. I create dummy variables for having been born in Natal, being born outside of Natal ('TCF'), and individual dummies for being born in the Transvaal, Cape, or Free State provinces. As noted above, I assume that individuals are educated in their province of birth and include municipality of birth fixed effects in all regressions.

Log income is defined as the natural log of reported yearly income for those who were employed at the time of the census. I find no selection into employment based on the policy change. A variable for years of education is obtained directly from the census.[15] There is good reason to worry that income is not perfectly reported in this dataset. However, this is a problem for the analysis only if misreporting differs across treatment and control provinces and treated and untreated cohorts so that it is correlated with treatment. Figures 1 to 4 do not seem to show that the spikes or dips in the average education or income levels by age differ substantially by birth province.

Finally, the census asks whether individuals can speak or can speak, read, and write, in their mother tongue as well as English and Afrikaans. I use these separate variables to construct an indicator variable for having some skill in English or Afrikaans, defined as being able to speak or speak, read, and write. I consider both languages together rather than individually due to the predominance of

15 Unfortunately, the 1980 census data lumps together grades 1 through 3 so anyone who attends school but does not pass third grade is assigned three years.

Table 4: Summary statistics for men, 1980

	All	Natal 'Old'	TCF 'Old'	Natal 'Young'	TCF 'Young'
Age	36.40	42.99	43.01	32.47	32.44
	(5.99)	(2.96)	(2.96)	(3.24)	(3.21)
Income	1415.53	1434.89	1300.21	1428.21	1320.22
	(1415.80)	(1576.18)	(1517.54)	(1527.59)	(1298.88)
Log(income)	7.043	7.066	7.002	7.089	7.053
	(0.808)	(0.821)	(0.835)	(0.771)	(0.796)
Years of school	4.24	3.49	3.71	4.30	4.64
	(3.64)	(3.61)	(3.61)	(3.66)	(3.62)
Percentage with no education	0.354	0.445	0.421	0.342	0.303
	(0.478)	(0.494)	(0.494)	(0.474)	(0.459)
Percentage with more than 4 years of education	0.487	0.387	0.424	0.486	0.540
	(0.494)	(0.487)	(0.494)	(0.499)	(0.498)
Average years of mother tongue education	3.57	2.85	2.27	4.00	4.34
	(3.01)	(2.68)	(1.91)	(3.25)	(3.24)
Speaks English or Afrikaans	0.615	0.363	0.633	0.429	0.685
	(0.486)	(0.481)	(0.481)	(0.495)	(0.47)
Literate	0.675	0.596	0.618	0.686	0.719
	(0.468)	(0.490)	(0.485)	(0.464)	(0.449)
Potential years of mother tongue exposure		6.00	4.00	8.00	8.00

Notes: Standard deviations in parentheses. The sample restricts to men between ages 28 and 48 who were born outside of homelands not enumerated in the 1980 census. The 'young' cohort is defined as being born after 1942 (or younger than 38 years old in 1980). Income and Log(Income) are reported for those with non-zero income.

Afrikaans in the Free State and Transvaal provinces and of English in the Natal and Cape provinces. Finally, I construct an indicator of literacy based on whether the individual can read and write in any language.[16]

Descriptive statistics for each cohort of men and women, respectively, are presented in Table 4 and Table 5. For both groups, education levels show an upward trend in both provinces, which appears to be partly driven by a larger

16 Results are not presented separately for reading and writing in only a native language because almost no individuals (0.4%) report reading and writing in English and Afrikaans but not in a native language.

Table 5: Summary statistics for women, 1980

	All	Natal 'Old'	TCF 'Old'	Natal 'Young'	TCF 'Young'
Age	36.52	43.02	43.05	32.48	32.46
	(6.04)	(2.99)	(2.98)	(3.27)	(3.26)
Income	694.26	706.03	661.66	731.86	706.09
	(1033.39)	(1022.45)	(1144.14)	(1085.75)	(949.72)
Log(Income)	6.230	6.319	6.156	6.353	6.245
	(0.939)	(0.853)	(0.957)	(0.868)	(0.946)
Years of school	3.73	3.23	3.17	3.92	4.10
	(3.57)	(3.44)	(3.49)	(3.55)	(3.60)
Percentage with no education	0.416	0.464	0.492	0.36	0.371
	(0.492)	(0.498)	(0.499)	(0.484)	(0.483)
Percentage with more than 4 years of education	0.438	0.368	0.373	0.449	0.487
	(0.496)	(0.483)	(0.483)	(0.497)	(0.50)
Average years of mother tongue education	3.55	2.60	2.10	3.85	4.40
	(3.08)	(2.67)	(1.94)	(3.21)	(3.28)
Speaks English or Afrikaans	0.520	0.323	0.517	0.381	0.588
	(0.499)	(0.467)	(0.499)	(0.485)	(0.492)
Literate	0.607	0.562	0.534	0.645	0.649
	(0.488)	(0.496)	(0.498)	(0.478)	(0.477)
Potential years of mother tongue exposure		6.00	4.00	8.00	8.00

Notes: Standard deviations in parentheses. The sample restricts to women between ages 28 and 48 who were born outside of homelands not enumerated in the 1980 census. The 'young' cohort is defined as being born after 1942 (or younger than 38 years old in 1980). Income and Log(Income) are reported for those with non-zero income.

percentage of individuals having attended school. Both literacy and the ability to speak English or Afrikaans are lower in Natal but increase between cohorts in both the treatment and control provinces.

Before analysing the effects of the language reform, I first verify that the policy change did induce a greater number of years of instruction in mother tongue. In particular, I expect that actual years of exposure increased more outside of Natal, where the years of potential exposure increased by four years, than in Natal where years of potential exposure increased by only two years. Indeed, I find that, for men, years of exposure increases by just over a year in Natal. Similar patterns hold for women. The increase in actual years of exposure does not equal the potential years of exposure because not every student stays in

school through grade six, seven, or eight. In fact, around half of the sample attained at least grade five so a significant part of my sample was affected by the reform. This is important given that my difference-in-difference coefficients are intent to treat estimates. To calculate a rough TOT estimate, I can divide my ITT estimate by the proportion treated, or about 0.5.

RESULTS

Difference-in-difference estimates

I begin by examining the effect of the language reform on wages and other adult outcomes for cohorts of men and women affected by the policy change in different ways across provinces. I consider the comparison between young (affected) and old cohorts and the difference between Natal and the other three provinces. In Table 6, difference in difference estimates of β for men in specification (1) are presented in Panel A for men and in Panel B for women. Separate regressions are presented for five different outcome variables. Recall that the coefficients represent the effect of increasing years of instruction in a student's mother tongue by two additional years.

Column (1) shows that, for men, the effect of increasing mother tongue instruction for two years was to raise wages later in life by 1.5%. This estimate is

Table 6: Difference in difference estimates of the effect of increasing mother tongue instruction by two years: men and women

	(1)	(2)	(3)	(4)	(5)
	Log Income	Years Education	White Ability	Literate	Employed
Panel A: Men					
TCF	0.015*	0.101***	−0.018*	0.010**	−0.002
	(0.005)	(0.006)	(0.006)	(0.002)	(0.006)
R-Squared	0.1074	0.1402	0.1831	0.1005	0.0564
N	1,204,076	1,581,750	1,581,750	1,581,750	1,581,750
Panel B: Women					
TCF	0.020**	0.181***	0.005	0.026***	0.007
	(0.005)	(0.013)	(0.005)	(0.002)	(0.009)
R-Squared	0.2179	0.2325	0.2754	0.1913	0.1229
N	612,240	1,627,532	1,627,532	1,627,532	1,627,532

Notes: Standard errors are clustered at the birth province level; p-values are calculated using the t-distribution with three degrees of freedom. Coefficients presented are the difference in difference coefficients (β): the coefficient on the interaction between the young cohort and the variable TCF. All regressions include fixed effects for municipality of birth and dummies for each age. The regressions restrict to ages 28–48 and to those born outside of homelands which were not enumerated in the 1980 census. *** p<0.01, ** p<0.05, * p<0.1

about 21% of the return to a year of education of 7%.[17] For women, the effect is slightly larger at 2.0%, but the difference across genders is not statistically significant. As noted above, a policy-maker might be more interested in knowing the actual effect on those who the policy affected, the TOT estimate. To recover the TOT effect, I can divide the ITT estimate by the proportion of the population who was affected. Because approximately 50% of my sample passed grade 5 or higher, this amounts to doubling the estimates above for a TOT estimate. In this case, the effect on the treated individuals in the sample was about a 3% increase in wages or 0.2 years of education for men; for women, it would be a 4% increase in wages of 0.36 years of education. These increases in wages are equivalent to about half a year of school so they are not fully accounted for by the increases in education levels.

The remainder of the table explores other outcomes which may help explain the increase in wages from this policy change. First, I consider whether the policy increased levels of education. For men, I find that there was a response to the policy change of about 0.1 years more of education. Slightly less than half of the effect of the program on wages can be explained by observed increases in years of education. That is, given a return to schooling of 7%, an increase of years of education by 0.1 would increase wages by 0.070% (=0.01*7*0.1) which is about half of the main effect of the program on wages. For women this effect was substantially larger at 0.18 years of education; education can account for just over 60% of the main effect on wages for women. The larger effect on women lends support to the hypothesis that mother tongue instruction increases accessibility to school and retention in school for girls. This also illustrates the fact that using education attainment in the difference-in-difference specification would be incorrect since education levels seem to have endogenously responded to the change.

Beyond years of schooling, two additional outcomes that may be affected directly by mother tongue education are ability to speak English or Afrikaans and the ability to read and write, either in one of these languages or in one's mother tongue. In column (3), for men I find negative effects of mother tongue instruction on the ability to speak English or Afrikaans ('White Ability'). At 1.8 percentage points, this estimate is quantitatively large, representing a decrease of 3%, on the national mean ability of 60%. There are no statistically significant effects on English or Afrikaans separately (results not shown), suggesting that the variation in the policy affected different languages differently. To examine the importance of English vs. Afrikaans, I split my sample in Table 7, which I discuss below. For women, there is no significant effect on the ability to speak a 'White' language.

Column (4) presents results for literacy in any language, measured as having the ability to speak, read, and write any language. The estimated effect of increasing mother tongue instruction by two years is a 1 percentage point increase

17 Author's own calculation from the 1980 census.

Table 7: Difference-in-difference effects for predominantly English-speaking birth municipalities, men and women

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Income	Years Education	White Ability	English Ability	Literate	Employed
Panel A: Men						
TCF	0.027**	0.098***	0.004**	0.007*	0.010***	−0.003
	(0.005)	(0.010)	(0.001)	(0.002)	(0.001)	(0.013)
R-Squared	0.0703	0.1486	0.1452	0.1327	0.0102	0.0304
N	532,948	767,798	767,798	767,798	767,798	767,798
Panel B: Women						
TCF	0.031**	0.140***	0.013**	0.012*	0.021***	0.004
	(0.009)	(0.023)	(0.003)	(0.004)	(0.001)	(0.008)
R-Squared	0.1697	0.2676	0.2411	0.2336	0.2224	0.1266
N	244,765	841,068	841,068	841,068	841,068	841,068

Notes: Standard errors are clustered at the birth province level; p-values are calculated using the t-distribution with three degrees of freedom. Coefficients presented are the difference in difference coefficients (β): the coefficient on the interaction between the young cohort and the variable TCF. All regressions include fixed effects for municipality of birth and dummies for each age. The regressions restrict to ages 28–48 and to those born outside of homelands which were not enumerated in the 1980 census. Additionally, they restrict to birth municipalities in which whites were predominantly English speaking (less than half of white men reported Afrikaans as their home language). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

in the probability of being literate in some language for men and an increase of 2.6 percentage points in the probability for women. Again, this effect is large, showing a 3.5% increase on the national mean of 41%.[18] This increase in literacy is indicative of positive human capital effects of the policy change. The policy seems to reduce spoken fluency in English or Afrikaans, which can be absorbed as a by-product of classroom instruction in English, but to have increased written literacy overall.

Finally, I consider whether higher wages might be a product of selection into employment induced by the policy change. I find no evidence that this is the case in Column (5). Estimates are close to zero and precisely estimated. Recall that the variable 'Employed' is only equal to one if the individual is employed and reports non-zero income; therefore there is no evidence that reporting zero income is different for those affected and unaffected by the policy.

18 Ideally, I would like to use this increase in literacy caused by the policy change for an IV estimate of the return to literacy in the 1980 South African labour market (as in Angrist & Lavy 1998). However, since the policy change also prompted a response in terms of education, it is not a valid instrument for literacy.

I next turn to the question of whether the effect of the policy was different in English- and Afrikaans-speaking areas. It is possible that the policy was more important in English speaking areas if individuals were more likely to need English for semi-skilled jobs. Because I do not know exactly which language was the primary language for each student, even at the birth municipality level, I use the 1980 census and classify birth municipalities based on the proportion of white men who were born there who report English or Afrikaans as their primary language. Because Afrikaans is more common than English, I use a cut-off of 0.4 where I assume the municipality is primarily Afrikaans if less than 40% of men speak English and that it speaks English if more than 40% of men speak English as a primary language.[19]

The main concern with splitting the sample into 'English' and 'Afrikaans' areas is that my control province, Natal, speaks mainly English (100 out of 118 birth municipalities are classified as English under my definition). Therefore, I present results only for birth municipalities that I classify as English-speaking. This amounts to comparing Natal to English-speaking municipalities in the other provinces.

Results are shown in Table 7. Coefficients are largely similar to Table 6. There are positive effects on wages of 2.7% and 3.1% for men and women, respectively. These numbers are about 50% larger than Table 6, providing some evidence that there were larger effects in English-speaking areas than the country as a whole. The effects on educational attainment are almost identical to those in Table 6. The interesting differences between English-speaking areas and the full country come with respect to the ability to speak English. There is an increase of 0.7 percentage points for men and 1.2 percentage points for women in the probability of speaking English. This suggests that the policy did have positive effects on English-speaking abilities in places where English was the primary labour market language. I calculated a 25% return to English speaking in the 1980 census, so this increase in English ability could help explain why the effect on wages was larger in this area.

Heterogeneity by province

I next allow for more flexibility between provinces and estimate a separate β for each province:

$$y_{ijm} = \alpha + \delta_1 young_j + \delta_2 transvaal_m + \delta_3 cape_m + \delta_4 freestate_m + \beta_1 young_j * transvaal_m + \beta_2 young_j * cape_m + \beta_3 young_j * freestate_m + \theta_j + \gamma_m + \varepsilon_{ijm} \quad (2)$$

where the same fixed effects are included as in specification (1). The parameters β_1, β_2, and β_3 represent the effect of increasing years of mother tongue

19 Results are similar with different cut-offs. In fact, results are stronger if I use a more stringent definition of 'English speaking' of more than 80% English speakers, but I lose a large part of my sample and the four provinces are arguably less comparable in this case.

Table 8: Alternative specifications: enter provinces separately, men

	(1)	(2)	(3)	(4)	(5)
	Log Income	Years Education	White ability	Literate Any	Employed
Cape	0.027***	0.098***	−0.012***	0.014***	−0.024***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Transvaal	0.009***	0.107***	−0.013***	0.008***	0.004***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
Free State	0.026***	0.082***	−0.041****	0.015***	−0.002
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
R-Squared	0.1024	0.1402	0.1832	0.1005	0.0565

Notes: N = 1,204,076 for column (1) and N=1,516,640 for the remaining columns. Standard errors are clustered at the birth province level; p-values are calculated using the t-distribution with three degrees of freedom. Coefficients presented are the difference in difference coefficients on the interaction between the young cohort and the respective province. All regressions include fixed effects for municipality of birth and dummies for each age. The regressions restrict to ages 28–48 and to men born outside of homelands which were not enumerated in 1980. *** p<0.01, ** p<0.05, * p<0.1

instruction by two years in the Transvaal, Cape, and Free State provinces, respectively.

Results are shown in Tables 8 and 9 for men and women, respectively. These tables present the estimates for each β_i in specification (3). For men, the effect on wages is higher in the Cape and Free State than in the Transvaal (2.7%, 2.6%

Table 9: Alternative specifications: enter provinces separately, women

	(1)	(2)	(3)	(4)	(5)
	Log Income	Years Education	White Ability	Literate Any	Employed
Cape	0.033***	0.228***	0.016***	0.031***	−0.009***
	(0.001)	(0.009)	(0.001)	(0.001)	(0.005)
Transvaal	0.022***	0.168***	0.008***	0.024***	0.016***
	(0.001)	(0.003)	(0.003)	(0.001)	(0.005)
Free State	0.002*	0.186***	−0.019***	0.030***	−0.014***
	(0.017)	(0.003)	(0.003)	(0.004)	(0.002)
R-Squared	0.2179	0.2320	0.2755	0.1913	0.1230

Notes: N = 643,236 in column (1) and N=1,695,944 for the remaining columns. Standard errors are clustered at the birth province level; p-values are calculated using the t-distribution with three degrees of freedom. Coefficients presented are the difference in difference coefficients on the interaction between the young cohort and the respective province. All regressions include fixed effects for municipality of birth and dummies for each age. The regressions restrict to ages 28–48 and to women born outside of homelands which were not enumerated in 1980. *** p<0.01, ** p<0.05, * p<0.1

versus 0.9%). The Transvaal has the highest average wages (see Figure 3), possibly due to the prevalence of jobs in mining. For education, the effect was between 0.08 and 0.1 years. In the Transvaal and Cape, students are 1.3 and 1.2 percentage points less likely to speak English or Afrikaans, respectively, while in the Free State they are 4.1 percentage points less likely. The effect on literacy is largest in the Cape and Free State at 1.4 and 1.5 percentage points, but smaller in the Transvaal at 0.8 percentage points. We see a significant effect on employment in the Cape Province, where the probability of being employed decreases 2.4 percentage points; if those who are not working are low skilled, we'd expect higher wages among the affected cohort.

In Table 9, the story is similar for women except with respect to wages. The largest effects on wages are seen in the Cape Province and Transvaal where wages increase by 3.3% and 2.2%, respectively. Female wages in the Free State increase only slightly and there seems to be a significant effect on employment here as well.

Robustness and threats to identification

In Table 10, I consider a variety of robustness checks. Column (1) shows the original results from Table 7. Column (2) looks only at individuals who were born in homelands which were not independent in 1980. The effect is larger in this group for men at 3.3%. For women, the effect is still an increase of 2.0%.

Table 10: Robustness. Outcome = Log(Income)

	(1)	(2)	(3)	(4)	(5)
	No Restriction	Restrict to those born in homelands	Allow students to be one year behind	Allow students to be two years behind	Age controls
Panel A: Men					
TCF	0.015*	0.033**	0.021*	0.021*	0.014*
	(0.005)	(0.006)	(0.007)	(0.008)	(0.005)
N	1,204,076	698,284	1,231,887	1,284,577	1,204,076
Panel B: Women					
TCF	0.020**	0.020*	0.021**	0.017*	0.019**
	(0.005)	(0.008)	(0.005)	(0.006)	(0.005)
N	612,240	170,678	627,241	655,629	612,240

Notes: Column (4) includes age and age-squared instead of a series of dummy variables for each age. Column (5) replaces birth municipality fixed effects with birth province fixed effects. Standard errors are clustered at the birth province level; p-values are calculated using the t-distribution with three degrees of freedom. Coefficients presented are the difference in difference coefficients (β) or the coefficient on the interaction between the young cohort and the respective province. Regressions include fixed effects for municipality of birth (columns 1–4) and dummies for each age. The regressions restrict to ages 28–48. *** p<0.01, ** p<0.05, * p<0.1

In columns (3) and (4), I show that results are robust to allowing students to be one year or two years behind where I predict them to be; that is, I change the cut-off to being born in 1943 or later and then 1944 or later. This is equivalent to assuming that the timing of the policy implementation was delayed one year or two years. In column (5), I show that results are robust to using age and age^2 as controls instead of a series of dummy variables.

One concern with the main identification strategy is that wages, education levels, and literacy are all increasing with birth year; younger cohorts have higher levels of education, wages and literacy. If these variables are increasing faster in the treated provinces, then this would bias the difference-in-difference results.

One option is to control for provincial-age trends, but then I cannot interpret my difference-in-difference coefficient as the causal treatment effect (Mora & Reggio 2014). I have also tried a triple-difference strategy using individuals who should not have been affected by the policy to control for any provincial-cohort trends. Results are similar (not shown) but it must be acknowledged that this strategy would hinge on the policy not using variation in schooling attainment which has already been shown to be endogenous to the policy change.

Another option is to control for white levels of education and income by cohort and birth municipality as a way of picking up any common trends that affect both groups; doing this in the regular regressions also does not change the results. Finally, one would be worried if we picked up a treatment effect after assigning 'treatment' to older cohorts and comparing the original untreated cohort to an even older cohort. I re-ran the same analysis as in specification (1) for cohorts aged 38 to 58 and assigned treatment to those aged 38 to 48. I see no statistically significant results and the coefficients are very close to zero, suggesting that, at least as far back as 10 years, there were no pre-trends in the treatment and control cohorts which could explain the results.

CONCLUSION

The Bantu Education Act provides a way to determine the effects of mother tongue instruction on a range of outcomes in 1980. The results taken together are consistent with the hypothesis that mother tongue instruction increased literacy and educational attainment. Mother tongue instruction appears to have had positive effects on school retention for both genders but especially for girls, which was one of the hypothesized effects.

This paper illustrates that the effects of mother-tongue instruction can be different depending on context. In Morocco, Angrist and Lavy (1998) found that Arabic instruction reduced French reading and writing skills and that this was reflected in lower wages. Angrist et al. (2008) find no effects of Spanish instruction on English skills. My results are entirely opposite in that mother-tongue instruction increased literacy and educational attainment, and this was reflected in higher labour market outcomes. My results are mostly, however, consistent

with Taylor and Coetzee's (2013) findings in favour of mother-tongue education in early years in contemporary South Africa.

Overall, it seems that mother-tongue instruction in this context might have had positive effects on school quality as evidenced through the increases in literacy of men and women. This cannot be disentangled from the relative benefits of teaching in English versus teaching English as a subject, as was done after the reform. The estimates presented in this paper most likely overstate the effects of simply changing the language of instruction, since they include effects on instructional quality and the ability of parents to help their children with homework.

While extrapolating results from apartheid South Africa to current policy in a very different country is misguided, this paper adds to our knowledge of an important time period in South African history and of the long-run effects of one of the components of the Bantu Education Act.

ACKNOWLEDGEMENTS

I am grateful to feedback from Leah Boustan, Dora Costa, Christian Dippel, Walker Hanlon, Carlos Flores, and seminar participants at the UCLA Applied Microeconomics and Economic History proseminars. The paper benefited substantially from comments from three anonymous referees.

REFERENCES

Akinnaso, F, 1993. Policy and experiment in mother tongue literacy in Nigeria. International Review of Education 39, 255–285.

Angrist, J & V Lavy, 1997. The effect of a change in language of instruction on the returns to schooling in Morocco. Journal of Labor Economics 15, 343–369.

Angrist, J, A Chin & R Godoy, 2008. Is Spanish-only schooling responsible for the Puerto Rican language gap? Journal of Development Economics 85, 105–128.

Benson, C, 2000. The primary bilingual education experiment in Mozambique: 1993 to 1997. International Journal of Bilingual Education and Bilingualism 3, 149–166.

Bertrand, M, E Duflo & S Mullainathan, 2004. How much should we trust difference-in-difference estimates? Quarterly Journal of Economics 119, 45–97.

Bleakley, H & A Chin, 2008. Language Skills and Earnings: Evidence from Childhood Immigrants. Review of Economics and Statistics 86, 481–496.

Cameron, AC & D Miller, 2014. A practitioner's guide to cluster-robust inference. Journal of Human Resources, forthcoming.

Cameron, A C, J Gelbach & D Miller, 2008. Bootstrap-based inference with clustered errors. Review of Economics and Statistics 90, 414–427.

Casale, D & D Posel, 2011. English language proficiency and earnings in a developing country: The case of South Africa. The Journal of Socio-Economics 40, 385–393.

Chiswick, B, 1991. Speaking, reading, and earnings among low-skilled immigrants. Journal of Labor Economics 9, 140–170.

Dustmann, C, 1994. Speaking Fluency, Writing Fluency, and the Earnings of Migrants. Discussion Paper no. 905. Center for Economic Policy Research.

Giliomee, H, 2009. A note on Bantu Education. South African Journal of Economics 77, 190–198.

Hartsthorne, K, 1953. Native Education in the Union of South Africa: A Summary of the Commission on Native Education in South Africa – U.G. 53–1951. South African Institute of Race Relations, Johannesburg, South Africa.

Harsthorne, K, 1992. Crisis and Challenge: Black Education 1910–1990. Oxford University Press, Cape Town.

Heugh, K. 2000. Giving good weight to multilingualism in South Africa. In Phillipson, R. (Ed) Rights to Language: Equity, Power, and Education. New Jersey: Lawrence Erlbaum, 122–159.

Lafon, M, 2009. The Impact of Language on Educational Access in South Africa. Create Pathways to Access Research Monograph No. 24. Educational Policy Unit, University of the Witswatersrand, South Africa.

Levinsohn, J, 2007. Globalization and the Returns to Speaking English in South Africa. NBER Working Paper 10985. National Bureau of Economic Research.

Mariotti, M, 2012. Labor markets during apartheid in South Africa. Economic History Review 65, 1100–1122.

McManus, W, W Gould & F Welch, 1983. Earnings of Hispanic men: The role of English language proficiency. Journal of Labor Economics 1, 101–103.

Mora, R & I Reggio, 2014. Treatment Effect Identification using Alternative Parallel Assumptions. UC3M Working Paper, Economics, No. 12–33.

Piper, B & E Miksic, 2011. The early grade reading assessment: Applications and interventions to improve early grade literacy. In Mother Tongue and Reading: Using Early Grade Reading Assessments to Investigate Language-of-Instruction Policy in East Africa. Research Triangle Park, NC: RTI Press, 139–182.

Soudien, C, 2002. Teachers' responses to the introduction of apartheid education. In Kallaway, Peter, (Ed), The History of Education under Apartheid. New York: Peter Lang, 202–245.

South African Bureau of Racial Affairs, 1955. Bantu Education: Oppression or Opportunity. Stellenbosch.

South African Data Archive Codebook to the 1980 Population Census. Available online: http://sada.nrf.ac.za/CodebookPDF/S0071.pdf (accessed 23 July 2014).

Taylor, N & P Vinjevold, 1999. Getting Learning Right. Report of the President Education Initiative Research Project, Joint Education Trust, Johannesburg.

Taylor, S & M Coetzee, 2013. Estimating the Impact of Language of Instruction in South African primary schools: A fixed effects approach. Stellenbosch Economic Working Papers 21/13, Stellenbosch University.