

Grade structure, educational attainment and labor market outcomes: Evidence from Botswana

Evan Borkum, Columbia University

September, 2009

Job Market Paper

Abstract: Many students in developing countries drop out of school at the junctures between different levels of education, yet little evidence exists on the extent to which changing the position of these junctures can affect educational attainment and other outcomes. This paper studies the impact of reforms to the position of the juncture between the lower and upper secondary levels in Botswana. I find that relocating the 10th year of education from upper to lower secondary, where access was much higher, induced an increase in educational attainment of 0.62 years for the group directly affected by this change. This was accompanied by an increase in labor force participation (3.9 percentage points) and a drop in unemployment (8.4 percentage points). Restricting to males in order to reduce sample selection bias, this group also selected into occupations with a higher skill level (0.19 of a standard deviation) and enjoyed higher wages (almost 16 percentage points). Consistent with these strong responses, IV estimates suggest high pecuniary returns of between 19.2 and 26.3 percentage points to the 10th year of education. Overall the results suggest that educational attainment responded strongly to an effective increase in access to the additional year of education, and that grade structure could have an important role to play in designing policies to increase educational attainment in developing countries.

JEL codes: I20, I28, J24

I would like to thank the Central Statistics Office of Botswana, particularly Anna Majelantle, Susan Matroos and Esther Wanjohi for assistance with obtaining the data. The generous assistance of the staff at the Education Statistics Unit (Ministry of Education) and the Botswana National Archives and Records Services is also much appreciated. I am grateful to Leigh Linden, Miguel Urquiola, Janet Currie and seminar participants at Columbia for helpful comments and advice. All errors are my own. Contact email: etb2102@columbia.edu.

1. Introduction

Cross-country data suggest a large amount of variation in the way in which the grades of an education system are configured into levels like primary and secondary (World Bank, 2009). Even amongst countries with the same total years of formal education, the allocation of these years between primary and secondary education tends to differ. In the case of secondary education, it is also common practice to distinguish between lower and upper secondary education, with the number of years allocated to each sublevel again exhibiting cross-country variation¹.

These distinctions are particularly important in developing countries, where the majority of individuals do not finish school. In many of these countries, enrollment bottlenecks are observed at the junctures between the various levels (primary, lower secondary and upper secondary). This could reflect that access is more limited in the next level (in terms of the supply of places), that costs are sharply higher and/or that students must pass a national exam for admission. These junctures therefore provide natural exit points from the formal education system. The national system of qualifications is usually structured accordingly, so that those who drop out at these points can enter the labor market or alternative vocational training with a qualification certifying their completion of a particular level.

One might therefore suspect that differences in grade structure in such countries could affect educational attainment and hence future outcomes by changing the position of these exit points. Despite the potential importance of this issue, existing studies on the effect of different grade structures are largely limited to analyzing the impact of the middle school structure in the U.S on student outcomes². However, since middle school grades are too early to be an exit point from formal education in the U.S, it is unclear

¹ For example, Indonesia offers 6 years of primary, 3 years of lower secondary and 3 years of upper secondary for a total of 12 years of school. On the other hand, while Zambia also offers a total of 12 years of school, this consists instead of 7 years of primary, 2 years of lower secondary and 3 years of upper secondary school.

² This middle school literature looks at outcomes such as test scores (Coldarci and Hancock, 2002; Rockoff and Lockwood, 2009), the incidence of disciplinary problems (Cook *et al.*, 2008), high school dropout rates (Alspaugh, 1998) and on-time high school completion (Bedard and Do, 2005).

whether these studies are especially relevant for policies aimed at improving educational attainment in the developing country context ³.

In order to provide some of the first evidence in this regard, I take advantage of changes in the structure of the secondary education system in Botswana. Botswana initially offered 3 years of lower secondary education and 2 years of upper secondary education, but reversed this structure in the late 1980's. A decade later, the structure reverted back to the original one. Since many individuals in Botswana tend to exit the schooling system after the completion of lower secondary education, these changes provide a useful natural experiment to investigate the impact of different grade configurations on educational attainment and labor market outcomes. The basic identification strategy is to compare individuals in cohorts that would have been subject to different structural regimes, being careful to control for other inter-cohort differences.

I find that educational attainment responded strongly to these changes in grade structure. Having the 10th year located in lower secondary rather than senior secondary led to an average increase in educational attainment of around 0.62 years for the students who tended to drop out from school around this point and were therefore directly affected by the change. The barrier to entering upper secondary school, driven in Botswana by a lack of access, thus appears to have been binding for many of these students. As a result, locating the 10th year of education in lower secondary – where access was higher – resulted in a large increase in average educational attainment.

Reduced form estimates suggest that the extension of lower secondary education by a year also had a large impact on labor market outcomes. For the affected group, the longer lower secondary regime was associated with an increase in labor force participation of 3.9 percentage points and a drop in unemployment of 8.4 percentage points. These gains were concentrated amongst females, even though the improvement in educational attainment was identical for both genders. Restricting to the male sample in order to limit possible sample selection bias, I find that the change in regime raised the occupational skill level of those employed by around 0.19 of a standard deviation and induced an increase of wages of almost 16 percentage points. I argue that finding these

³ Outside of this literature, Oosterbeek and Webbing (2007) study the extension of tertiary vocational education in Holland from 3 to 4 years. However, this does not address any of the issues raised above for pre-tertiary education, particularly in developing countries.

strong labor market effects in this context implies that the 10th year of education had considerable human capital value. Finally, I use the nature of the structural regime as an instrument to estimate pecuniary returns to the 10th year of education, which range from 19.2 to 26.3 percentage points. These high returns are consistent with the strong response to the change in access to the 10th year of education associated with the changes in structural regime.

This study contributes to the broader literature on government policies that can influence educational attainment and hence future labor market outcomes such as wages. Several studies have focused on policies that improve enrollment, which may eventually translate into increased attainment, such as fee reductions (see for example Borkum, 2009 and Barrera-Osorio, Linden and Urquiola, 2007) and Conditional Cash Transfers (see for example Schultz, 2004). A policy closely related to this paper is an expansion of access to education simply by building more schools. Duflo (2001) investigates a large primary school construction project in Indonesia, and finds substantial increases in educational attainment and wages for cohorts affected by the expansion. The issue of access is highly relevant here since it is one of the main reasons that students in Botswana do not proceed to upper secondary education. Extending the duration of lower secondary education by a year can therefore be viewed as an increase in access to the additional year.

Perhaps the most common policy to boost educational attainment that has been considered to date in the literature is that of compulsory schooling laws. Angrist and Krueger (1991) were the first to use features of compulsory schooling laws in the U.S. to investigate returns to education, by noting that these laws would force students born later in the year to stay in school longer. Subsequent studies have taken advantage of mandated changes in compulsory schooling laws across a range of developed countries⁴. Most of these studies agree that compulsory schooling laws are effective in increasing educational attainment, although there is some variability in the extent to which this affects later outcomes for those affected by these laws. While changes in compulsory schooling laws are in the same spirit as this paper in that they reflect a national policy that can induce changes in educational attainment, they are less relevant from a policy

⁴ Examples include Harmon and Walker (1995) for the U.K., Acemoglu and Angrist (2000) for the U.S., Oreopoulos (2006) for the U.S., Canada and the U.K., and Pischke and von Wachter (2008) for Germany.

perspective in developing countries because of the difficulty of enforcement. Indeed, while certain developing countries have already enacted compulsory schooling legislation, this legislation often exists only on paper. On the other hand, changes in the grade structure of the national education system might induce changes in educational attainment by altering the nature of the institutional environment faced by parents and students.

The remainder of this paper is structured as follows. Section 2 provides some background by describing the relevant policy reforms in detail and illustrating their implementation. Section 3 discusses the identification strategy while section 4 describes the data used for estimation. Section 5 investigates the impact of the reforms on educational attainment. Section 6 presents reduced form evidence for the impact of the policy reforms on a variety of labor market outcomes and section 7 estimates returns to the 10th year of education using an IV approach. Section 8 concludes.

2. Background and first stage

a. Background: Reforms to the grade structure of secondary education

Formal education in Botswana is divided into three distinct levels: primary, junior (lower) secondary and senior (upper) secondary school. As in many developing countries, national exams are taken at the end of each level. Since access to higher levels is limited in the sense that demand exceeds supply, places are rationed based primarily on the results in these exams. Students who pass receive a certificate with which they can enter the labor market or vocational training if they do not progress to the next level. At the time of independence from Britain in 1966, the structure of the education system was referred to as “7+3+2”: 7 years of primary, 3 years of junior secondary and 2 years of senior secondary school.

The first National Commission on Education (NCE) convened in 1977, at a time when relatively few students continued their formal education beyond the primary level. This was mainly due to a lack of access to secondary school⁵, and resulted in a large number of individuals leaving school at a young age with a primary school certificate that

⁵ In 1977 only 20% of primary school graduates could be accommodated in government secondary schools (a further 15% went into community-run “unaided” schools).

offered limited opportunities in the labor market or for further training. The Commission viewed this situation with concern, and recommended a long-term target of 9 years of universal basic education. While this would obviously require a substantial increase in access to junior secondary school, the Commission also recommended that the structure of the education system be changed from 7+3+2 to 6+3+3. With this new grade structure it was envisaged that all students would complete the first two levels (primary and junior secondary school) and thus attain 9 years of education, with the best graduates then admitted into senior secondary school on a competitive basis.

However, there was a strong argument against making this change in grade structure immediately. The main problem was that initiating this change before junior secondary access had been improved would result in many students leaving school after grade 6⁶, with only 6 years of primary education rather than the current 7. Moreover, the junior secondary system was not believed to be capable of handling the double cohort that would enter junior secondary in the year of the change. A transition via a temporary 7+2+3 structure was suggested as a solution, with the final switch to 6+3+3 deferred until junior secondary access had been improved. Crucially for this paper, the temporary change from a 7+3+2 to a 7+2+3 structure was implemented for the cohort entering grade 8 in 1986⁷.

A second NCE that met in 1992/1993 brought up many problems associated with the temporary 7+2+3 structure. In particular, the Commission's report mentions a widespread perception that 2 years of junior secondary education was insufficient to prepare students for work or further training and also did not offer sufficient time for them to adjust from primary to secondary school. There was thus a strong feeling that a structure involving 3 years of junior secondary education should be reintroduced as soon as possible.

⁶ In Botswana primary school grades are called "standards" (1-7) and secondary school grades are called "forms" (I-V). However, to avoid potential confusion I will use the term "grade" throughout. In my terminology the schooling system therefore runs from grade 1 to grade 12, with grades 1-7 comprising primary school and grades 8-12 comprising secondary school.

⁷ Regardless of the structure, the examinations leading to the Junior Certificate (the next qualification after the Primary School Leaving Exam) always took place at the terminal grade of junior secondary school. So when the structure changed from 7+3+2 to 7+2+3, the Junior Certificate exams switched from being held in grade 10 (testing a 3 year curriculum) to being held in grade 9 (testing a 2 year curriculum).

However, completing the change in structure to 6+3+3 was no longer viewed as desirable, due to the outstanding logistical difficulty of a double cohort entering junior secondary school and the problem of transferring teachers with different qualifications and salary scales from primary to secondary schools. Instead, a switch back to the original 7+3+2 structure was recommended. In order to maintain current levels of access, this was to be made possible by the construction of extra classrooms in existing junior secondary schools and the hiring of additional junior secondary teachers. Implementation of the change in structure was for the cohort entering grade 8 in 1996 – exactly a decade after the cohort experiencing the initial reform. In summary, the grade structure actually implemented in Botswana involved an initial 7+3+2 structure which was switched to 7+2+3 for a period of a decade and then switched back.

b. First stage: Implementation of reforms

As a country with a small population concentrated in a few major centers, Botswana's education system is highly centralized and oversees a relatively small number of schools. It is therefore a useful setting in which to study policy reforms of this nature. First, the small number of schools means that policy changes can be implemented sharply. For example, there were only around 200 junior secondary schools in existence at the time of the switch back to the 7+3+2 system in 1998. The extra construction and teacher hiring required to accommodate an extra year in junior secondary could therefore be completed in a short period of time so that the reform could be implemented across the entire system without delay. Second, monitoring of the implementation of policy changes by the national Ministry of Education is relatively straightforward since officials need only follow up with a small number of schools that are largely concentrated in a few major centers.

Administrative data from the annual *Education Statistics Reports* (Central Statistics Office of Botswana) can be used to directly illustrate the reforms to the structure of secondary education. Figure 1 shows enrollment in each grade of secondary school. The enrollment level in grade 10 – which is the grade that switches between junior and secondary school – initially follows that in grades 8 and 9 (the junior secondary grades). In 1988 there is a steep drop in grade 10 enrollment, which then

follows the enrollment level in grades 11 and 12 (the senior secondary grades). The difference in access between the two levels is also clearly illustrated, with enrollment in senior secondary grades much lower than that in junior secondary grades⁸. In 1998, grade 10 enrollment jumps dramatically to once again follow enrollment in grades 8 and 9.

This pattern is consistent with the sharp changes in the structure of the secondary education system described earlier. That is, grade 10 was initially part of the junior secondary school system so that access and hence enrollment was similar to that in the other junior secondary grades. It was then switched to senior secondary school in 1988 (for the cohort entering grade 8 in 1986), where lower access meant that enrollment dropped to the level of the other senior secondary grades until it was finally switched back to junior secondary school in 1998 (for the cohort entering grade 8 in 1996).

3. Identification

The ideal dataset would include information on the educational cohort of each individual in the sample, since it would then be clear whether that individual experienced a 9 year or a 10 year junior secondary regime⁹. Unfortunately, all the data used in this paper only include information on current age (and hence year of birth). In the absence of a binding rule on the age of school entry, this is insufficient to determine educational cohort. A typical approach in such a situation is to assume that all individuals enter school at the modal age of entry, as in (for example) Angrist and Lavy (1997).

However, I choose to adopt an alternative approach in order to reduce the extent of measurement error in the treatment assignment. This approach uses administrative data on the age distribution for each educational cohort to calculate the probability that each birth cohort experienced a 10 year regime. Specifically, I use Bayes' rule which states that:

$$\Pr(edcohort | yob) \propto \Pr(yob | edcohort) * \Pr(edcohort)$$

⁸ Enrollment levels give an accurate picture of the level of access in this context because the available places in each level of secondary school are filled to capacity.

⁹ Although the duration of junior secondary school would have been 2 or 3 years it is more convenient to refer to the total number of years of education including 7 years of primary school. Therefore I refer to a 9 year or 10 year regime throughout.

where *edcohort* is educational cohort and *yob* is year of birth. Using administrative data on the age of entry from the annual *Education Statistics Reports*, I can obtain $\Pr(\text{yob}|\text{edcohort})$ for all educational cohorts from 1966 (Botswana independence) onwards¹⁰. I then use grade 1 enrollment for each educational cohort to approximate $\Pr(\text{edcohort})$ for the population and apply Bayes' rule. The $\Pr(\text{edcohort}|\text{yob})$ measure is then normalized and yields the probability of being in a 10 year regime by noting that educational cohorts entering grade 8 before 1986 or after 1995 (or alternatively entering grade 1 before 1979 or after 1988) would have experienced a 10 year regime¹¹.

These calculations lead to a U-shaped pattern based on year of birth, as illustrated in Figure 2. For comparison, Figure 2 also illustrates the treatment assignment assuming that individuals enter school at the modal age¹². The U-shape in Figure 2 based on the Bayes' rule construction shows that those born prior to 1970 would almost certainly have experienced a 10 year regime. The probability of experiencing a 10 year regime then gradually declines to zero, so that those born in 1974 are virtually certain of having experienced a 9 year regime. It then recovers again so that those born after 1982 are certain of having experienced the switch back to the 10 year regime.

I can then estimate regressions of the form:

$$Y_{ijt} = \alpha + \beta.\text{prcohort}10_j + \pi.X_{ijt} + \theta.Z_j + t_j + \mu_t + \varepsilon_{ijt} \quad (1)$$

¹⁰ It is clear from the data that there is significant variation in the age of entry and that there have been large changes in its distribution over time even though the modal age is 7 years virtually throughout (see appendix Figure A1). Moreover, it is worth noting that although the distribution in Figure A1 does shift over time towards entry at older ages, there is little visual evidence of a strong systematic response in the age of entry corresponding to the two changes in regime. This helps to rule out that the reduced form effects identified are simply due to an age of entry effect.

¹¹ Of course this assumes no repetition, which would add to any measurement error in the treatment assignment. Unfortunately, the data do not contain information on grade repetition for individual respondents. In the aggregate, automatic promotion in primary schools was introduced after independence but repetition in grade 4 was allowed from 1981 and remained at 10-15% thereafter. Perhaps more importantly, there was substantial repetition in grade 7 (the terminal grade of primary school), since many students took the Primary School Leaving Exam more than once in order to improve their scores and gain admission to junior secondary school. The repetition rate in grade 7 peaked at 30% in 1983 and then declined steadily to under 5% a decade later. Section b of the appendix argues that repetition considerations do not have a large effect on the reduced form estimates here.

¹² Section a of the appendix shows that the reliability ratio, defined as the fraction of the observed variability that is distinguishable from measurement error, is roughly 0.22 higher using the Bayes' rule method as opposed to the modal age method. Besides considering the issue of grade repetition, section b of the appendix also compares the main results in the paper using these two alternative methods of constructing the instrument.

where Y_{ijt} is some outcome variable for individual i in birth cohort j observed in survey year t and $prcohort10_j$ is the probability of experiencing a 10 year regime. X_{ijt} is a vector of individual level controls such as gender and age, Z_j is a vector of birth cohort controls, t_j is a linear birth cohort trend, μ_t is a survey year dummy and ε_{ijt} is the usual error term. The coefficient of interest is β , which identifies the reduced form impact of experiencing a 10 year regime on the relevant outcome. In all the regressions I cluster standard errors by birth cohort in order to account for possible intra-cohort correlation.

The major concern for identification arises from the fact that $prcohort10_j$ is assigned based on birth cohort. This variable simply reflects the probability that individuals in a particular birth cohort entered junior secondary school in years other than the 1986-1995 period when the 9 year regime was in effect. If individuals in birth cohorts entering junior secondary in the period corresponding to a particular type of structural regime share some characteristic that affects the outcome variable through an unrelated channel, then the estimate of β is biased.

One strength of this study that helps to mitigate such a concern is that there were two sharp changes in structural regime – the initial change from 7+3+2 to 7+2+3 and the switch back a decade later. Therefore the identification does not rest merely on the worsening or improvement of outcomes for cohorts born after a certain period, which could be a result of many factors. Instead, if the changes in grade structure are really driving the estimates for a certain outcome, then the cohort pattern for that outcome should display a worsening (or improvement) of outcomes followed by a reversal that corresponds to the timing of the changes in regime. In other words one would expect that outcome to display a U-shaped (or inverse U-shaped) pattern that corresponds to the probability of experiencing a 10 year regime as illustrated in Figure 2. In subsequent sections I will present graphical evidence that this is indeed the case for educational attainment and several labor market outcomes.

Although it is difficult to imagine other contemporaneous changes that would generate such patterns, I nevertheless attempt to control for important differences at the

birth cohort level which are denoted Z_j in equation (1). The most obvious and potentially important confounding changes over the period of study involve changes in the education system and the economy as a whole and are illustrated in the four panels of Figure 3. The unit for the horizontal axis in each case is educational cohort, denoted by the year of entry into junior secondary school while the vertical lines demarcate the timing of the two changes in structural regime. Educational cohorts between the two vertical lines are those subject to the 9 year regime¹³.

Panel A of Figure 3 illustrates changes in access to secondary education by plotting progression rates into junior and senior secondary school¹⁴. It is evident that there was a steady and dramatic increase in the progression rate to junior secondary that begins at around the same time as the initial change in regime. This corresponds to a substantial increase in access to junior secondary education as a result of the large-scale construction of Community Junior Secondary Schools by the government. These schools offered only junior secondary grades and were attended predominantly on a day-scholar basis. They were introduced as an alternative to the existing model of expanding access by constructing large secondary schools with boarding facilities which offered all secondary grades. The progression rate to senior secondary school initially decreases as the junior secondary expansion gains momentum since senior secondary access could not keep up, but eventually improves for later educational cohorts.

Controlling for some measure of these progression rates is important for two reasons. First, increased access may be reflected in the unobserved ability of individuals. For example, since access to junior secondary increased for cohorts in the 9 year regime, these cohorts are likely to consist of lower ability individuals on average and experience worse outcomes compared to those in the first 10 year regime (although not compared to those in the second) regardless of the grade structure. Second, there may be a pure supply

¹³ Although the characteristics plotted in Figure 3 vary at the educational cohort level, all those in the same birth cohort are assigned the same probability distribution over educational cohorts. Therefore the j subscript in equation (1) refers to birth cohort. In the regressions, these controls are assigned probabilistically. That is, I use the normalized $\Pr(edcohort | yob)$ derived from Bayes' rule as weights in calculating a weighted average of the relevant control for each individual based on their birth cohort.

¹⁴ For junior secondary school the progression rate is the ratio of grade 8 enrollment to grade 7 enrollment in the previous year (i.e. for the same educational cohort). For senior secondary school it is similarly defined as the ratio of enrollment in the grade of entry (grade 10 or grade 11 depending on the structure in that year) to the preceding grade's enrollment in the previous year.

side effect due to the changes in access, which is particularly relevant for labor market outcomes. For example, the larger number of individuals entering the labor market with a junior secondary education in the 9 year regime could depress wages for this group regardless of the change in structure. Panel B of Figure 3 shows the size of the cohort leaving school at the junior secondary level (i.e. between grade 8 and the terminal grade of junior secondary school) and illustrates a sharp increase in this cohort size measure beginning at around the time of the first change in regime which then flattens out. Since the cohort size measure in panel B is closely related to the changes in access illustrated in panel A, the former is the variable I choose to include in the analysis as a parsimonious control for all these potential effects.

Panel C of Figure 3 illustrates the measure that I use to control for school quality, namely the secondary school student-teacher ratio. The ratio for each educational cohort is calculated for the year in which it enters junior secondary school (grade 8). The student-teacher ratio is initially flat but begins a steady decline at around the time of the first change in regime, suggesting an improvement in quality along this limited dimension. While this variable is certainly not an exhaustive control for school quality, the plot in panel C suggests that along this dimension school quality did not change in a systematic way with the changes in regime. Nevertheless, I include it as a control in some of the specifications.

Finally, although survey year fixed effects account for labor market conditions in the year in which individuals are observed, different cohorts could have experienced different labor market conditions at the time of their entry into the labor market. These conditions may have had long-term effects on their future outcomes. Accordingly, panel D of Figure 3 plots the growth rate of real GDP in the year in which each educational cohort would have entered the labor market, calculating this year by assuming that the full junior secondary course was completed¹⁵. The average growth rate is high: Botswana enjoyed one of the world's highest sustained growth rates in the post-independence era, due largely to the discovery of vast diamond deposits. However, the plot suggests a

¹⁵ In the regressions I use the actual reported educational attainment for each individual to calculate the year of labor market entry. Strictly speaking, this variable therefore belongs in X_{ijt} rather than Z_j , but it is more useful conceptually to think of it as a cohort level difference.

marked slowing in the growth rate corresponding to the first change in regime that then recovers somewhat by the time of the second change. This pattern suggests that individuals in the 9 year regime would, on average, have experienced worse initial labor market conditions using this measure compared to those in the first 10 year regime (although again not compared to those in the second). This control is therefore important in ruling out that these worsening initial labor market conditions corresponding to the 9 year regime are driving the estimates.

4. Data and summary statistics

Four main datasets are available for the analysis in this study. The first two are the decennial censuses of 1991 and 2001, which are full censuses of the population. The second two are the inter-census Labor Force Surveys (LFS's) of 1995/96 and 2005/06. These are stratified two stage probability samples designed to be representative of the full population when the relevant sample weights are applied. All datasets contain information on educational attainment and on respondents' ages, which I use to assign birth cohort. The LFS data also include detailed labor force outcomes such as employment status, occupation, hours worked and wages.

For each dataset, I restrict the sample to individuals who are at least 20 years old and born after 1960. The restriction to those aged 20 and above is in order to allow individuals to completely resolve their educational decisions, while the restriction to those born after 1960 is to avoid dealing with individuals who entered school prior to independence in 1966 when little information is available on the educational system. By pooling the datasets I obtain a repeated cross section consisting of four surveys separated by 5 year intervals and spanning the period 1991-2006, with the labor force outcomes only available in the two LFS's¹⁶. Panel A of Table 1 presents basic demographic summary statistics for the full sample broken down by education level. Column 1 is for all individuals reporting non-zero educational attainment. Columns 2-4 look at individuals with educational attainment corresponding to the completion of different levels of the educational system: 7 years (primary school), 9 or 10 years (junior

¹⁶ The full repeated cross section consists of: the 1991 census with individuals aged 20-30, the 1995/96 LFS with individuals aged 20-35, the 2001 census with individuals aged 20-40 and the 2005/06 LFS with individuals aged 20-45.

secondary school) and 12 years (senior secondary school) respectively. The average individual in the sample is in their late 20's, while the gender ratio is slightly skewed towards females but less so for those with a full 12 years of education (column 4). Average educational attainment is around 8.8 years for those reporting any formal education (column 1).

Besides educational attainment, I also investigate the impact of the change in structure on several labor market outcomes which are available in the LFS's. The first outcomes that I focus on are labor force participation and unemployment¹⁷. I also investigate occupation-related outcomes using the ISCO-88 classification of occupational groups that allows one to assign a skill level from 1-4 to each occupation¹⁸. These outcomes are binary variables for a particular skill level and the actual skill level itself, normalized to represent standard deviations from the sample mean of those reporting non-zero educational attainment. These occupational measures are only available for those currently employed. Finally, I investigate the log hourly wage (monthly cash receipts plus benefits divided by monthly hours) conditional on reporting a positive wage.

The summary statistics for these labor market outcomes in panel B of Table 1 show that, while labor force participation is similar across various educational groups, the other outcomes are more favorable for those with higher educational attainment. That is, those with more education are less likely to be unemployed and tend to work in higher skilled occupations and earn higher wages conditional on employment. On average senior secondary graduates (12 years of education) in column 4 earn almost five times as much as primary school graduates (7 years of education) in column 2.

¹⁷ I follow the Central Statistics Office of Botswana's definitions for the construction of these variables. The unemployment dummy takes value 0 if an individual did any work in the previous week either as paid employment (cash or kind) or as self employment for profit or family gain. It takes value 1 for individuals who were not employed in the reference period but were available for work. The labor force participation dummy takes value 1 if an individual was either employed or unemployed in the reference period and 0 otherwise. Individuals not in the labor force include those doing solely unpaid domestic work in their own house, studying or not working because they were sick or retired. The more standard definition of these variables involves reclassifying individuals who were available for work but not actively seeking it as out of the labor force. Using this alternative definition leads to quantitatively similar results.

¹⁸ Occupations in skill level 1 are "elementary occupations" such as construction workers, cattle herders and cleaners. Common examples of skill level 2 occupations are shop salespeople, bricklayers and office clerks, while for skill level 3 examples include nurses, primary school teachers and technicians. Skill level 4 consists of occupations such as post-primary teaching professionals, accountants and engineers.

Finally, as discussed, I use auxiliary data from the annual *Education Statistics Reports* to introduce controls for important differences across birth cohorts. This is in order to deal with confounding cohort-level factors that may affect outcomes but are not directly related to the changes in structure. These data are from an annual survey of schools, include detailed information on schools, enrollments and teachers and were obtained for most of the years since independence. This allows for a fairly complete picture of the evolution of the education system over time and for the construction of cohort level controls for access and quality. These administrative data also provide the age distributions in grade 1 that are used to construct the probability of experiencing a 10 year regime as described in the previous section.

5. Reduced form impact of structural regime on educational attainment

a. Graphical evidence

The central premise of this paper is that students are more likely to drop out from school at the transition between educational levels. This should be reflected in large peaks in the distribution of educational attainment corresponding to the terminal grades of primary school (7 years), junior secondary school (9 or 10 years depending on the regime) and senior secondary school (12 years). Figure 4 displays a histogram of educational attainment using the full pooled sample of those reporting non-zero years of education. Since the terminal grade for junior secondary school differs according to the regime, I ignore those for whom the nature of the regime is ambiguous (i.e. those with $prcohort10_j$ not equal to zero or unity in Figure 2) and plot the histogram separately for those who would have experienced a 9 year and 10 year regime with certainty.

Regardless of the nature of the structural regime, there are large peaks at 7 and 12 years of education, corresponding to the completion of primary and senior secondary school respectively. The other large peak varies by the nature of the regime: it is located at 9 years for those in a 9 year regime ($prcohort10_j = 0$) and at 10 years for those in a 10 year regime ($prcohort10_j = 1$). This corresponds to the terminal grade of junior secondary school. The figure therefore shows that the vast majority of individuals tends to exit the educational system at the points corresponding to the junctures between levels

and suggests that Botswana is indeed an appropriate setting in which to study the impact of reforms to the position of these junctures.

In order to investigate whether the changes in structure were effective in changing educational attainment, it is necessary to introduce a time dimension. If these changes did have such an effect, one would expect to see changes in the proportion of individuals attaining 9 versus 10 years of education that correspond to the timing of the changes in regime. Figure 5 plots the proportion of individuals reporting each level of educational attainment by birth cohort for the full pooled sample, conditional on reporting some secondary education (i.e. more than 7 years). The figure shows that there are large differences in the proportions reporting 9 or 10 years of education by birth cohort. Those born prior to the late 1960's (who definitely experienced the 10 year structure) were far more likely to attain 10 years of education, even though some individuals (10-15% of the sample) were leaving school with 9 years of education. However, for older birth cohorts the proportion attaining 10 years of education becomes U-shaped and is mirrored by an inverse U-shape for the proportion attaining 9 years of education. At the same time, the proportion attaining 8 or 11 years of education remains very low while that attaining 12 years increases gradually in a more linear fashion. Figure 6 addresses the issue of educational attainment more directly by plotting the mean educational attainment by birth cohort. For the full sample (again conditional on reporting some secondary education), mean educational attainment again displays a broadly U-shaped pattern. This pattern is even starker when one considers only those reporting 9 or 10 years of education, for whom the change in the duration of junior secondary schooling should have had the largest effect.

These U-shaped patterns are consistent with a story in which the changing structural regime affected the educational attainment for those who tended to leave school after completing their junior secondary education and would therefore have attained an additional year in the longer regime. The change is not quite symmetric: Figure 5 suggests that when the structure reverted back to a 10 regime, fewer students tended to drop out after 9 years compared to during the earlier 10 year regime. In particular, comparing the fraction of individuals with 9 years of education in Figure 5 at either end of the domain suggests that around 10-15% of individuals who definitely experienced the

initial 10 year regime attained 9 years of education, while only around 5% of those who experienced the later 10 year regime did so. This is likely due to a trend towards improved efficiency in the education system over time.

One should also note that the timing of the changes in the distribution does not correspond exactly to the probabilities in Figure 2 although the general pattern is consistent. For example, the decline in the proportion with 10 years of education and increase in the proportion with 9 years is well underway for the 1969 birth cohort in Figures 5 and 6, although according to Figure 2 this birth cohort was almost certain to have experienced a 10 year regime. This discrepancy is consistent with some degree of grade repetition which, as mentioned earlier, leads to measurement error in the treatment assignment. Nevertheless, there is still strong graphical evidence that suggests that the change in grade structure did indeed have an impact on educational attainment.

b. Reduced form estimates

Panel A of Table 2 estimates the impact on educational attainment by regressing years of education on $prcohort10_j$ and several control variables for the full pooled sample reporting at least some secondary education, as in equation (1). As mentioned, since $prcohort10_j$ is assigned at the birth cohort level, it is important to control for other factors that differ systematically by cohort. Column 1 shows the simple regression of years of education on $prcohort10_j$ and a set of survey year dummies with no other control variables. Column 2 adds the size of the cohort dropping out at junior secondary level in order to control for different levels of access, together with the student-teacher ratio in the year of entry into junior secondary school in order to control for differences in educational quality. This specification also includes a linear birth cohort trend to account for unrelated systematic inter-cohort changes in educational attainment. Finally, column 3 investigates possible asymmetry in the impact of the initial and later changes in structural regime by adding the term $prcohort10_j * I_{>1973}$, the interaction between the probability of experiencing a 10 year regime and a dummy for being in a post-1973 birth cohort (and hence experiencing the later 10 year regime if at all).

The estimates in columns 1 of panel A are unaffected by the addition of controls in column 2 and suggest an increase of around 0.32 years of education associated with a switch from the 9 year to the 10 year regime. Column 3 provides little econometric evidence for the asymmetry evident in Figure 6 once the various controls are included. It should be noted that the adjusted R^2 in these regressions is rather low so that one can only explain under 3% of the variation in years of education for this sample.

The estimates in panel B of Table 2 are from similar regressions, only restricting the sample to those reporting 9 or 10 years of education. This is the sample for which one would expect the changes in regime to be binding and therefore to have the largest effect. For this sample, the point estimate in column 1 is around 0.62. Adding the controls for access and quality and the linear cohort trend in column 2 again has virtually no effect. The estimate in column 2 is quite precise so that we can reject any change lower than 0.53 of a year at the 95% confidence level. In addition, the adjusted R^2 for these regressions is around 30% and over – far larger than for the full secondary school sample. This sample, for which the change in regime has much stronger explanatory power, is the focus of the subsequent reduced form estimates for labor market outcomes.

The estimates in column 3 of panel B provide some evidence of asymmetry between the initial and later changes in regime for the sample with 9 or 10 years of education. The point estimates suggest that the initial switch to a 9 year regime led to a decrease of around 0.53 years of education, while the switch back to a 10 year regime led to an increase of around 0.74 years (the sum of the coefficients), although the coefficient on the interaction term is only significant at the 10% level¹⁹. This asymmetry again reflects the fact that the dropout rate within junior secondary school was lower for those experiencing the later 10 year regime compared to the earlier one, as illustrated in Figure 5. Therefore the later change was binding on more individuals, and would have had a stronger impact on educational attainment.

¹⁹ To see the asymmetric impact of the two changes in regime more directly, one can rewrite equation (1) augmented with $\gamma.prcohort10 * I_{>1973}$ as:

$$Y_{ijt} = \alpha - \beta.(1 - prcohort10_j + prcohort10_j * I_{>1973}) + \delta.prcohort10_j * I_{>1973} + \varepsilon_{ijt}$$

where controls have been omitted for clarity and $\delta = \gamma + \beta$. One can then show that the coefficient $-\beta$ yields the impact of the initial reduction from 10 to 9 years, which is -0.528 with a standard error of 0.064. The coefficient δ directly yields the impact of the later extension from 9 to 10 years, which is 0.742 with a standard error of 0.069.

As a check of whether the effects on educational attainment are indeed concentrated amongst those with 9 or 10 years of education, Table 3 runs regressions similar to those in panel A of Table 2, only using a dummy for having completed each year of education as the dependent variable in a set of linear probability models. The results indeed suggest that the impact of changing from a 9 year to a 10 year regime arises almost exclusively due to an increase of around 34% (in the full specification of column 2) in the probability of completing the 10th year of education. This is again consistent with there being a large group of students dropping out of school at the juncture between junior and senior secondary education, so that the change in regime is binding for them in terms of their educational attainment. In contrast, there is little impact on the probability of completing senior secondary school (12 years). This is despite the fact that the additional costs of proceeding to complete senior secondary are higher in the 9 year regime because they entail an additional year of education.

The evidence thus suggests that shifting the barrier between junior and senior secondary school by a year had a strong impact on educational attainment. The estimates imply that experiencing a 10 year as opposed to a 9 year regime is associated with an average increase in educational attainment of around 0.62 years for the group directly affected i.e. those who tended to leave school after 9 or 10 years of education. Given the probable measurement error remaining in the *prcohort10_j* variable, this estimate should be interpreted as a lower bound for the true increase in educational attainment. Section 6 investigates how the changes in structural regime affected various labor market outcomes.

c. Policy implications: grade structure and access

In many developing countries, access to education (by which I mean the extent to which demand exceeds supply) reflects the grade structure of the education system. In Botswana access decreases sharply as one moves from junior to senior secondary due to the much lower supply of secondary schools, with places rationed based on exam results at the end of junior secondary. As discussed, the barrier imposed by this lack of access results in many students dropping out of school after the completion of junior secondary. The response to the changes in grade structure in Botswana can therefore be interpreted as a

response to changes in access to the 10th year of education. Access was higher when this 10th year was located in junior secondary instead of senior secondary and therefore more individuals completed it.

More generally, allowing for a change in grade structure may give policymakers more flexibility as they seek to improve access to various levels of education – a key element in improving educational attainment. The standard policy to improve access is simply to construct more schools, as in Duflo’s (2001) study in Indonesia. However, such programs can be very costly. For example, dramatically improving access by constructing more three-year senior secondary schools in Botswana would have been prohibitively expensive, since the upper grades in these schools require specialized facilities (such as science labs) and teachers with a higher level of training. Instead, access to the 10th year of education was improved rapidly and fairly inexpensively by moving it to junior secondary school and simply constructing an additional classroom and hiring an additional teacher or two in each school in order to accommodate the additional year. Construction of more (two-year) senior secondary schools then took place at a much more gradual pace in subsequent years.

Another practical example of a case in which a change in grade structure could be used to improve access is a country that has “combined” secondary schools with no distinction between lower and upper secondary. Making such a distinction between levels and constructing new schools offering only lower secondary grades is likely to be a far cheaper and more feasible way of increasing access than constructing more of these combined schools. These examples illustrate how considering grade structure as a policy tool can lead to a wider range of policy options compared to the traditional policy of expanding access by simply constructing more schools taking the grade structure as given.

6. Reduced form impact of structural regime on labor market outcomes

Since the changes in regime had a large effect on educational attainment, particularly for the group with 9 or 10 years of education, one might expect them to have had an impact on labor market outcomes too. The first two available outcomes in the data are labor force participation and unemployment. Then, conditional on employment, one can

investigate the reduced form effect of the changes in regime on occupational choice using the four ISCO-88 occupational skill categories and the normalized skill level. Finally, one can investigate the impact on the hourly wage (which I express in logs) and break this down into an effect due to occupational choice and a within-occupation effect.

a. Labor force participation and unemployment

The solid markers in Figure 7 are mean residuals for the labor force participation and unemployment outcomes by birth cohort after controlling for a cubic in age and a survey year dummy, with a fifth order polynomial (solid line) fitted in order to aid the visual impression. These residuals represent the variation in outcomes across birth cohorts that cannot be explained by a simple flexible age trend, controlling for survey year effects in order to justify pooling data across surveys. The sample is restricted to those reporting 9 or 10 years of education: as discussed above, this is the sample for which the changes in structural regime would have been likely to be binding. For reference, I include the pattern of mean educational attainment by birth cohort (hollow markers and dashed line) which is reproduced from Figure 6.

The plots for both labor force participation and unemployment display a similar pattern in that they begin to decrease (labor force participation) or increase (unemployment) for late 1960's birth cohorts. They then adopt a U-shaped (labor force participation) or inverse U-shaped (unemployment) pattern for 1970's and 1980's birth cohorts that corresponds broadly to the changes in educational attainment. The fact that these outcomes are U-shaped is encouraging since it is difficult to explain such a pattern by other contemporaneous differences between birth cohorts. Nevertheless, in the econometric analysis for these and other outcomes I will attempt to control for some of these potentially confounding cohort-level differences, as discussed earlier.

Table 4 provides the econometric evidence for labor force participation and unemployment from regressions of the form of equation (1). The regressions are probits and report the marginal effect on $prcohort10_j$ at the sample mean values of the covariates. The basic specification shown in column 1 involves regressing the outcome of interest on $prcohort10_j$, demographic controls (male dummy, married dummy and

Gaborone dummy), a cubic in age, a survey year dummy and a linear birth cohort trend²⁰. Column 2 reproduces the same specification, only using a quartic rather than a cubic polynomial in age. Column 3 reverts to the cubic age polynomial, and adds controls for other potentially confounding cohort-level differences. These are the number of individuals in the cohort leaving school at junior secondary level (a control for changes in access and related cohort size effects), the secondary school student-teacher ratio experienced in grade 8 (a control for school quality) and the rate of GDP growth in the year of labor market entry (a control for initial labor market conditions).

Columns 4 and 5 break down the results by gender, restricting the sample to females and males respectively. It is important to point out here that the effects of the structural regime on educational attainment are virtually identical by gender (results not shown). This is not necessarily surprising since there is not a large difference in junior secondary enrollment rates by gender in Botswana so that both genders might be expected to respond in a similar manner to a change in regime in terms of educational attainment. Therefore, a direct comparison of the reduced form labor market effects by gender is a sensible one.

The estimate for labor force participation in columns 1-3 of Table 4 is fairly insensitive to the specification. Column 3 is the preferred specification, and the point estimate there suggests an average increase in labor force participation of 3.9 percentage points associated with a change from a 9 year to a 10 year regime (4.3% of the mean)²¹. The corresponding estimate for unemployment is reduced slightly by the addition of cohort level controls, but the point estimate in column 3 still suggests a highly significant decrease in unemployment of 8.4 percentage points (23.1% of the mean).

²⁰ If the surveys were each held in a single calendar year, then the survey year dummy and the linear age term would imply a linear birth cohort trend. Since the LFS's are held over two calendar years this is not quite true, although the collinearity is so high that adding the explicit linear birth cohort trend makes little difference in practice. In addition, since the treatment is at the birth cohort level and the identification is from variation across birth cohorts, I do not include higher order birth cohort trends for fear of spuriously soaking up variation caused by the treatment. Indeed, a sufficiently high order birth cohort trend would account for almost all of the variation in outcomes across birth cohorts. As a robustness check, I did include a quadratic birth cohort trend in the reduced form regressions for labor market outcomes: none of the estimates changed substantially with the exception of the estimate for labor force participation which decreased in magnitude and significance.

²¹ Since most of the estimates are relatively insensitive to the order of the age polynomial, the specification of column 3 (with the full set of controls and a simple cubic age polynomial) is the preferred specification.

Columns 4 and 5 suggest substantial heterogeneity in these effects by gender. The positive estimate for labor force participation is driven mainly by females, for whom it increased by 5.0 percentage points. In regressions not shown, it appears that this increase was driven by a decrease in the fraction of females engaged in household chores. Similarly, the estimate for unemployment is much higher in magnitude and significance for females: a drop of 10.2 percentage points (significant at the 1% level) compared to 4.1 percentage points for males (significant only at the 10% level).

b. Occupational skill and wages

The above results suggest that the analogous reduced form estimates for occupational skill and wages on the full sample could be subject to sample selection bias. Consider the occupational skill variable, which is only available for individuals currently employed. Since it was shown above that the nature of the structural regime affects the probability of employment, it also affects the probability of observing the occupational skill variable. As a result, there is a potential sample selection problem. For example, assume that only low “ability” individuals are unemployed if subject to the 9 year regime. When the regime switches to 10 years, these marginal individuals enter employment, but are likely to be unemployed in low skill occupations due to their low ability. The presence of these individuals in the data for the 10 year but not the 9 year regime would therefore cause a downwards bias in the estimated impact of the longer structural regime on occupational skill²². A similar argument holds for the log wage variable, since this is also only available for those currently employed²³.

Since the impact of the change in structure on male unemployment and hence the potential for selection bias is far lower than for the full sample, the following estimates

²² Since unemployment status is only observed for those participating in the labor force, one may be concerned that the earlier unemployment estimates for the full sample are also subject to sample selection bias. However, since the rate of labor force participation is around 90 percent, the scope for sample selection bias here is low.

²³ Actually, log wages are only reported for a subset of those currently employed, namely those individuals working as paid employees (the main categories of employment for which this is not the case are those working in subsistence agriculture and the self-employed). However, one can show that the structural regime has no effect on the probability of reporting a wage conditional on employment. Therefore, any sample selection bias for log wages is likely occurring through the impact of the structural regime on employment status itself.

are restricted to males²⁴. Section c of the appendix discusses the issue of sample selection bias in greater detail and argues that, if anything, it is indeed likely to be downwards so that the estimated effects can be interpreted as lower bounds.

Accordingly, Figure 8 displays a plot of residuals for the normalized occupational skill measure that is analogous to the plots for labor force participation and unemployment in Figure 7, only restricted to the male sample. The plot displays a similar U-shaped pattern to before. Again this suggests that the switch to a 9 year regime resulted in a drop in overall skill for those with 9 or 10 years of education that was reversed when the regime switched back to 10 years. Table 5 presents the econometric evidence on occupational skill using similar specifications to those in Table 4. The outcomes investigated here are dummies for each of the four occupational skill levels together with the overall normalized occupational skill measure plotted in Figure 7.

The estimates for the occupational skill dummies in Table 5 suggest that the strongest net effect of the switch to a 10 year regime was to increase the probability of being in an occupation of skill level 3. This estimate is reasonably stable across specifications and is 5.5 percentage points in the preferred specification of column 3, significant at the 1% level. The probabilities of being in occupational skill levels 1 or 2 each display a decrease that is comparable in magnitude but lower in statistical significance to the effect for skill level 3. The estimate for occupational skill using the normalized measure is 0.19 standard deviations in the preferred specification, again highly statistically significant.

For log wages, the residuals plotted in Figure 9 for the male sample exhibit the same familiar U-shaped pattern as the other labor market outcomes, although the smaller sample size leads to a somewhat more nebulous pattern. The econometric analysis for log wages is in Table 6, where the basic specification in column 1 suggests that wages are 13.4 percent higher for cohorts in a 10 year regime. This point estimate is slightly sensitive to the addition of controls, with the preferred specification of column 3

²⁴ For the sake of completeness, I report the results for females in Table A3 of the appendix. The impact on female labor market outcomes is generally stronger than for males, but it is unclear to what extent this is driven by differential sample selection bias.

suggesting that wages are around 15.8 percent higher for those experiencing a 10 year regime. Again these estimates are all significant at the 1% level.

Table 7 investigates how much of the difference in wages is attributable to occupational choice using the preferred specification of column 3 in Table 6 (the results are qualitatively similar if one uses one of the other specifications). Column 1 reproduces the initial estimate, only restricted to those reporting occupational skill in order to keep the same sample throughout the subsequent exercise. This restriction has little impact on the estimate although it does reduce its statistical significance slightly. Including the normalized occupational skill level in this regression (column 2) reduces the coefficient on $prcohort10_j$ by around 27%. Controlling for occupation using the full set of occupational skill dummies instead (column 3) reduces the original estimate by around 41% but also renders it statistically insignificant. Overall, these results suggest that the estimated effect of the change in regime on wages is attributable both to the change in occupational choice and to an increase in wages within occupational skill levels.

c. Interpretation: what are the reduced form effects capturing?

One concern for interpretation is that the two structural regimes could have differed in other ways besides the number of years comprising each level of education that are not controlled for in the regressions. This would make it difficult to attribute the estimated reduced form effects solely to the impact of grade structure. A particular concern is that of quality, by which I mean both the level of resources allocated to the education system and the content of education. While I have attempted to control for changes in resource allocation using the teacher-student ratio, this is a fairly crude control that may not adequately reflect other differences in resources such as teacher quality that could have differed systematically by the type of regime. With regards to content, several studies in the U.S. have suggested that the nature of the courses taken in school may be important for later labor market outcomes (see for example Rose and Betts, 2004 and Goodman, 2008). Evidence from the report of the 1992/1993 NCE suggests that both changes in structural regime were indeed accompanied by changes to the junior secondary curriculum in terms of the types and number of courses offered.

However, while the education system certainly did experience other changes along these dimensions, there is little evidence that these changes were systematically related to the changes in structure. The fact that there were two changes in regime again proves useful in this regard, since it allows one to disentangle unidirectional trends over time from the changes in structure. The only other measure of quality available (although not in all years) is the proportion of untrained teachers, which exhibits a general decline over time. This trend is much like that for the student-teacher ratio and points to a general tendency towards improved quality over time. Similarly, the changes in content documented in the NCE report were largely in the same direction over time. This involved a movement towards a less academically focused, more practical curriculum and the broadening of course offerings to include technical and commercial subjects. Therefore, while I cannot entirely rule out the possibility that other features of the two types of structural regime are playing some role in the reduced form estimates, this role is likely to be a minor one relative to the change in grade structure.

d. Interpretation: signaling and human capital

In interpreting the results in a broad theoretical framework, several facts regarding the institutional setting in Botswana are worth reemphasizing. First, admission into higher levels of education (junior and senior secondary) is based almost entirely on the results in the exam taken at the completion of the previous level. Essentially there is a cutoff score for entry each year that is adjusted according to the number of slots available, with the vast majority of individuals accepting a slot if offered one. Second, the formal national certificates of achievement that are recognized in the labor market are limited to the qualifications obtained at the end of each level of education. Finally, as evident from Figure 4, we observe stacking at the junctures between levels and few individuals dropping out within levels regardless of the structural regime. Most individuals with 9 or 10 years of education would therefore have dropped out at the same point in the system, namely at the end of junior secondary school, and would therefore have the same qualification (a junior secondary certificate).

The situation in Botswana can thus be crudely characterized as one in which individuals are separated into different levels of schooling based on their “ability” due to

the rationing of slots at higher levels based on exam results, with the group in each level attaining the full number of years prescribed by the structural regime. This pattern of educational attainment can be generated by a simple signaling model along the lines of Spence (1973) in which workers are paid their expected productivity and additional schooling has no human capital value. In such a setting, employers infer productivity and pay wages based on observable formal qualifications. Provided that the costs of completing an additional year of junior secondary are relatively low, individuals who are admitted into junior secondary will complete the full junior secondary course regardless of the structural regime in order to obtain a qualification of signaling value and will thus attain an extra year of education in the longer regime. This is indeed what is observed in the data.

However, such a model also implies that junior secondary graduates in the two regimes should earn the same in the labor market. This is because junior secondary graduates in both regimes possess a qualification that suggests they are sufficiently able to enter junior secondary but not sufficiently able to enter senior secondary. In other words, conditional on having the same level of access in both regimes (and recall that access is controlled for in the regressions), the nature of the signal and hence wages for junior secondary graduates should be the same. This aspect of the pure signaling framework is clearly at odds with the previous empirical results.

Adapting the basic signaling framework to include human capital yields more realistic predictions. If schooling contributes to the production function (but again employers only observe formal qualifications) then one can obtain the same results as before regarding educational attainment. Provided that the costs of an additional year are not too high, individuals admitted into junior secondary will again respond to a longer regime by attaining an extra year of education in order to obtain the junior secondary qualification. In this case, however, individuals with the same qualification but different years of education are not identical. The additional year makes junior certificate holders with 10 years of education more productive than those with 9 years and, since employers can infer (at least crudely) which regime an individual experienced, the former are rewarded with higher wages. The empirical results therefore argue against a pure

signaling model and suggest that the 10th year of education also has human capital value in the labor market²⁵.

7. Returns to the 10th year of education

The nature of the structural regime can be used as an instrument to estimate the pecuniary returns to the 10th year of education in Botswana. The IV point estimate of returns to the 10th year of education for males can be obtained by simply scaling up the reduced form estimate for log wages for the group with 9 or 10 years of education by the average change in educational attainment for the same group. The range in point estimates for log wages is 13.4 to 15.8 percentage points for this group depending on the specification (Table 6) and the estimate for educational attainment is around 0.62 (Table 2). The estimate of returns to the 10th year of education obtained in this manner thus ranges from 21.6 to 25.5 percentage points.

More formally, Table 8 presents estimates for the returns to the 10th year of education using a variety of specifications that are the same as those used for estimating the reduced form effect on log wages in Table 6. Panel A presents the OLS estimates of these returns by regressing log wages on educational attainment for the male sample with 9 or 10 years of education. Panel B presents analogous IV estimates using $prcohort10_j$ as an instrument for educational attainment, estimated by two-stage least squares (2SLS).

There are three main reasons why OLS and IV estimates of returns to education can differ, assuming that the instrument is valid. First, the OLS estimates may be contaminated by selection bias in that those who choose to complete an additional year education may be different along some unobservable dimension that also results in them earning higher wages for unrelated reasons (e.g. “ability”). Second, as pointed out by Card (1999) and others, the IV and OLS estimates may differ because they refer to different populations. In particular, one might expect the IV estimates to be larger because they are Local Average Treatment Effects (LATE's) for the population who can

²⁵ One could obtain similar predictions in a pure human capital model in which each individual's optimal educational choice has a corner solution that shifts as the change in regime alters her choice set. Such a model seems far less natural in the setting of Botswana than one based on signaling given the strong signal of ability resulting from selective admission into higher levels and the heavy reliance on formal qualifications in the labor market. Nevertheless, it cannot be entirely ruled out: the empirical evidence only allows one to reject a pure signaling model.

be induced to switch from 9 to 10 years of education, which may have higher marginal returns to an extra year of education than the entire group with 9 or 10 years of education. Finally, there may be measurement error in reported educational attainment which, if this error is classical in nature, causes OLS estimates of the returns to education to be attenuated.

In the current setting, concerns about selection bias and the interpretation of LATE are likely to be relatively minor. This is because, as discussed earlier, individuals are generally separated into different levels of schooling based on exam results, with the vast majority individuals in each level attaining the full number of years prescribed by the structural regime. Conditional on the level of access, most individuals who would have selected 9 or 10 years of education should therefore not be too dissimilar in average unobservable characteristics since they would have attained their level of education due to the exogenous structure of the education system, rather than due to sorting. Therefore the effects of selection bias are likely to be small, while the group induced to attain an extra year of education should also largely overlap with the group as a whole, as in Oreopoulos (2006). *A priori* one would therefore expect the IV estimates to be similar to the OLS estimates, or possibly smaller if measurement error is present.

The OLS estimate in the first column of panel A of Table 8 indicates that the 10th year of education is associated with an increase in wages of 17.5 percentage points. Adding cohort-level controls increases this estimate to 23.1 percentage points in the specification of column 3. The OLS point estimates here are slightly larger than those of Simphambe (2000), who estimates pecuniary returns to education in Botswana (with a different dataset) at around 16 percent using a basic Mincer-type specification that is linear in years of education. However, Siphambe also points out that a linear specification is probably inappropriate due to vastly different returns by education level, and estimates returns of 83 percent to the entire (3 year) junior secondary cycle using a more flexible specification.

The IV estimates are presented in panel B of Table 8. Again they are somewhat sensitive to the addition of cohort level controls, but are quite close to the corresponding OLS estimates within the same column. The range of IV point estimates is 19.2 to 26.3

percentage points, similar to that expected from the simple rescaling exercise performed initially²⁶.

8. Conclusion

The results in this paper suggest that the grade structure of the secondary education system in Botswana played an important role in determining educational attainment. Shifting the 10th year of education from the senior secondary to the junior secondary level, where access was higher, induced an increase in educational attainment of around 0.62 years for the group directly affected by this change. This increase in the duration of junior secondary education also had large reduced form effects on labor market outcomes for the affected group. Cohorts experiencing the longer junior secondary regime enjoyed significantly higher labor force participation and lower unemployment, selected into occupations requiring a higher skill level and benefited from higher wages.

The strong response of educational attainment to a longer junior secondary regime seems reasonable in Botswana since opportunity costs for those without a junior secondary qualification are relatively low, as are direct junior secondary schooling costs. Therefore one might have expected a strong response as individuals admitted into junior secondary decided that the labor market benefits of the extra year of education outweighed the additional costs. Indeed, it appears that the rationing of places imposed a binding constraint on most individuals' educational choices so that they tended to complete the full number of years in the highest level to which they gained admission.

A caveat is that one should exercise caution in applying these results directly to other settings where, for example, marginal schooling costs are high. If this had been the case in Botswana, the switch to a longer regime may have led to some individuals dropping out earlier as the benefits of completing junior secondary were no longer sufficient to compensate for the large costs of the additional year. However, it seems likely that in most settings the costs of an additional year or two of education within the same level should be relatively low so that a similar response might be expected.

²⁶ The slight difference is due to the fact that the reduced form estimates for educational attainment and log wages do not use the exact same covariates. Therefore, computing the IV estimate for log wages by scaling up is not technically correct – although in practice this makes little difference compared to implementing 2SLS.

From the more general perspective of designing policies to improve educational attainment in developing countries, these results suggest that grade structure is a neglected margin of the policy debate. Instead of being viewed as a given constraint (often inherited from the colonial era), it should rather be viewed as a deliberate policy tool. This is particularly true for countries such as Botswana in which access to higher levels of education is limited and is constraining individuals' educational attainment. An expansion of access this context may be prohibitively costly if it takes the pre-existing grade structure as given, while altering the grade structure may make such an expansion feasible. In this way Botswana was able to rapidly increase access to the 10th year of education relatively inexpensively by reverting to a three-year junior secondary structure.

There is in fact a historical precedent from the developed world for using grade structure to improve educational attainment. Goldin and Katz (2008) document that most American elementary schools in the early 20th century terminated at grade 8, while access to public high school was still limited. As a result, many individuals tended to drop out after completing grade 8. In order to encourage these individuals to stay in school for an additional year, reformers in some districts established public junior high schools that covered grades 7-9. These junior high schools became widespread in just a few years, and contributed to the increase in educational attainment in the US over this period.

References

- Acemoglu, D. and Angrist, J. (2000); "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws"; in B. Bernanke and K. Rogoff (Eds.); *NBER Macroeconomics Annual 2000*; Cambridge, MIT Press, 2001.
- Alspaugh, J. W. (1998); "The Relationship of School to School Transitions and School Size to High School Dropout Rates"; *The High School Journal*; 81: 154–160.
- Angrist, J. (1997); "Conditional Independence in Sample Selection Models"; *Economics Letters*; 54: 103-112.
- Angrist, J. and Krueger, A. (1991); "Does Compulsory Schooling Attendance Affect Schooling and Earnings?"; *Quarterly Journal of Economics*; 106: 979-1014.
- Angrist, J. and Lavy, V. (1997); "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco"; *Journal of Labor Economics*; 15(1): S48-76.
- Barrera-Osorio, F., Linden, L. and Urquiola, M. (2007), "The Effects of User Fee Reductions on Enrollment: Evidence from a Quasi-Experiment"; Columbia University Working Paper.
- Bedard, K., and Do, C. (2005); "Are Middle Schools more Effective? The Impact of School Structure on Student Outcomes"; *Journal of Human Resources*; 40: 660–682.
- Borkum, E. (2009); "Can Eliminating School Fees in Poor Districts Boost Enrollment? Evidence from South Africa"; Columbia University Working Paper.
- Card, D. (1999); "The Causal Effect of Education on Earnings"; In: Ashenfelter, O. and Card, D. (Eds) *Handbook of Labor Economics* Volume 3; Amsterdam: Elsevier.
- Central Statistics Office of Botswana; *Education Statistics Reports* (various years); Government of Botswana Press, Gaborone.
- Coldarci, T. and Hancock, J. (2002); "Grade-span Configuration"; *Journal of Research in Rural Education*; 17(3): 189-192.
- Cook, P.J., MacCoun, R., Muschkin, C. and Vigdor, J. (2008); "The Negative Impacts of Starting Middle School in Sixth Grade"; *Journal of Policy Analysis and Management*; 27(1): 104-121.
- Duflo, E. (2001); "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*; 91(4): 795–813.
- Goodman, J. (2008); "The Labor of Division: Returns to Compulsory Mathematics Coursework"; Columbia University Working Paper.
- Goldin, C. and Katz, L. (2008); *The Race Between Education and Technology*; Belknap Press of Harvard University Press: Cambridge, MA.
- Harmon, C. and Walker, I. (1995); "Estimates of the Economic Return to Schooling for the United Kingdom"; *American Economic Review*; 85(5): 1278-86.
- Heckman, J. (1979); "Sample Selection Bias as a Specification Error"; *Econometrica*; 47(1): 153-161.
- Horowitz, J.L. and Manski, C.F. (2000); "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data"; *Journal of the American Statistical Association*; 95: 77-84.
- Krueger, A. and Lindahl, M. (2001); "Education for Growth: Why and for Whom?"; *Journal of Economic Literature*; 39: 1101-1136.

- Lee, D. (2008); "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects"; *Review of Economic Studies*; 76(3): 1071-1102
- Oosterbeek, H. and Webbing, D. (2007); "Wage Effects of an Extra Year of Basic Vocational Education"; *Economics of Education Review*; 26: 408-419.
- Oreopoulos, P. (2006); "Estimating Average and Local Average Treatment Effects of Education when Compulsory School Laws Really Matter "; *American Economic Review*; 96(1): 152-175.
- Pischke, J. and von Wachter, T. (2008); "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation"; *Review of Economics and Statistics*; 90(3): 592-598.
- Rockoff, J.E. and Lockwood, B.B. (2009); "Stuck in the Middle: Educational Impacts of Grade Configuration"; Columbia Business School Working Paper.
- Rose, H. and Betts, J.R. (2004); "The Effect of High School Courses on Earnings"; *Review of Economics and Statistics*; 86(2): 497-513.
- Schultz, T.P. (2004); "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program"; *Journal of Development Economics*; 74(1):199-250.
- Siphambe, H.K. (2000); "Rates of return to education in Botswana"; *Economics of Education Review*; 19(3): 291-300.
- Spence, M.A. (1973); "Job Market Signaling"; *Quarterly Journal of Economics*; 87: 355-374.
- World Bank (2009); EdStats Database.
- World Bank (2008); World Development Indicators.

Appendix

a. Reliability ratio for different constructions of the instrument

Following Krueger and Lindahl (2001), note that we have two possible measures of the treatment both measured with (let us assume) classical error: $T_1 = T^* + e_1$ (Bayes' rule method) and $T_2 = T^* + e_2$ (modal age method). The reliability ratio for T_k , where $k \in \{1, 2\}$, is defined as:

$$RR_k = \frac{\text{var}(T^*)}{\text{var}(T^*) + \text{var}(e_k)} \quad (\text{A1})$$

If e_1 and e_2 are uncorrelated, then the estimate of the reliability ratio given by:

$$\hat{RR}_k = \frac{\text{cov}(T_1, T_2)}{\text{var}(T_k)} \quad (\text{A2})$$

converges in probability to (A1). This can be estimated by a simple binary regression of one measure on the other. However, in the current context e_1 and e_2 are likely to be positively correlated since the two measures are derived using the same birth cohort information for the same individuals. As a result, the estimate of the reliability ratio given by (A2) is biased upwards by the amount $\frac{\text{cov}(e_1, e_2)}{\text{var}(T_k)}$.

Although this makes it difficult to directly estimate (A1), it turns out that $\text{var}(T_1) \approx \text{var}(T_2)$. Therefore, the bias is similar for both measures and we have a result on the difference between the two reliability ratios, namely that $RR_1 - RR_2$ can be approximated by $\hat{RR}_1 - \hat{RR}_2$, which yields a point estimate of 0.22 with a negligible standard error. In other words, using the Bayes' rule method results in an instrument that has roughly 22 percentage points more signal than the modal age method.

b. Robustness to different constructions of the instrument

The probability of experiencing a 10 year regime, the instrument here, was constructed by applying Bayes' rule using administrative information on the age of school entry as described in section 3. In this section, I investigate the sensitivity of the results to variations on this method of constructing the instrument. The first modification is to

allow for the probability of grade repetition, since a comparison of Figure 2 and Figures 5 and 6 suggests that the assigned probabilities do not quite match the timing of the changes in educational attainment. Assuming that each individual would only have repeated once at most, repetition would affect the value of the instrument for individuals entering grade 1 in 1978 and 1988. These are the last educational cohorts experiencing the first and second structural regimes respectively, under the assumption of no repetition.

As a baseline, column 1 of Table A1 reproduces the earlier reduced form estimates for the sample with 9 or 10 years of education using the preferred specification in each case. Column 2 repeats these estimates, with a slight modification of the instrument to take into account repetition. To construct the instrument in this case, Bayes' rule is applied as before to obtain a probability distribution for each individual over educational cohorts. Cohorts entering grade 1 before 1978 or after 1988 are again assumed to experience a 10 year regime. The only difference from before is that cohorts entering in 1978 and 1988 are now assumed to experience a 10 year regime with probability 0.5. In other words, individuals are assumed to have a 50% probability of having repeated a grade once prior to entering junior secondary school, which affects the instrument for boundary cohorts. Column 3 takes this assumption further by assigning these boundary cohorts to a 9 and 10 year regime respectively i.e. it assumes a 100% probability of having repeated a grade once prior to entering junior secondary school. The estimates in column 2 are fairly similar in magnitude to those in column 1 for all of the variables, although in general slightly higher. Even with the extreme and unrealistic repetition assumption in column 3, the estimates are not substantially affected.

An alternative way to assign the instrument is to simply assume that each individual enters school at the modal age, which is 7 years throughout virtually the entire post-independence period. The instrument is then a binary variable that takes value 1 if an individual would have entered grade 1 before 1979 or after 1988 (assuming no repetition). Given the large variation in the age of entry (see Figure A1), one might expect that this would introduce substantial measurement error compared to the more careful method that assigns the instrument probabilistically. Indeed, as discussed in section a of this appendix, one can show that the reliability ratio of this measure is around

22 percentage points lower than the Bayes' rule measure. Column 4 of Table A1 repeats the estimates in column 1 using this alternative construction. These estimates are always lower in magnitude than the corresponding estimates in column 1. The estimate for educational attainment is around 12% lower (panel A), while the estimates for the various labor market outcomes (panel B) are between 13% and 18% lower. Column 5 again uses the modal age to assign the instrument but assumes a 50% chance of repetition, while column 6 assumes that everyone repeats a year for sure (this is equivalent to assuming that everyone starts school at age 8 without any repetition). The estimates again do not change dramatically, although the proportional change is somewhat higher compared to the introduction of repetition using the Bayes' rule measure in columns 1-3.

The evidence here is largely consistent with the estimates based on the assumption that individuals enter school at the modal age leading to greater measurement error and hence to attenuation bias in the reduced form estimates compared to those using Bayes' rule. This suggests that the latter method, which is used throughout the paper, is preferred. Using this method, the estimates are not substantially affected by assumptions on grade repetition, even under the extreme and unrealistic assumption that all individuals repeat a grade.

c. Bounds to address sample selection bias

As discussed in section 6, the issue of sample selection bias poses a potential problem for the reduced form estimates for occupational skill and log wages (and, to a lesser extent, for unemployment). In order to mitigate this problem, the estimates for occupational skill and log wages were restricted to the male sample for which the extent of sample selection bias is likely to be lower. This section attempts to deal more carefully with sample selection bias by placing further bounds on the estimates presented earlier.

There are essentially two methods available for dealing with sample selection bias of the nature discussed in this paper. The first method is to explicitly model the selection process and correct for selection bias by including a selection term in the outcome equation (see for example Heckman, 1979 and Angrist, 1997). However, for this approach to be of any practical use it is necessary to have some variables that affect the

probability of selection but not the outcome equation. In the current setting and with the available data, this does not seem plausible.

The second method aims to impose non-parametric bounds on the treatment effect by comparing “best case” and “worst case” scenarios for the selection pattern. For example, Horowitz and Manski’s (2000) approach essentially involves imputing missing values using the highest and lowest values in the sample. However, this method tends to lead to very wide bounds. Instead I adopt Lee’s (2008) approach, which instead of imputing values simply trims the upper or lower tail of the distribution of the outcome variable and attributes the remaining distribution to inframarginal individuals who are not subject to selection.

Specifically, Lee shows that one can obtain bounds on a binary treatment effect, T , in the presence of sample selection bias by computing the treatment effect in two cases:

Case 1: The top p % of outcomes for the treated group ($T = 1$) are from marginal individuals induced to select by the treatment.

Case 2; The bottom p % of outcomes for the treated group ($T = 1$) are from marginal individuals induced to select by the treatment.

In each case, the distribution of the outcome variable conditional on $T = 1$ is trimmed at the critical percentile, p . These marginal individuals are effectively ignored and the treatment effect recalculated, yielding a lower and upper bound for the treatment effect (cases 1 and 2 respectively). The critical p for trimming the conditional distribution of the outcome variable is given by:

$$p = \frac{\Pr[Y_{observed} | T = 1] - \Pr[Y_{observed} | T = 0]}{\Pr[Y_{observed} | T = 1]}$$

where $Y_{observed}$ means that the outcome variable is not missing. A key assumption here is monotonicity i.e. that treatment can only affect sample selection in one direction.

There two main issues in applying this method directly in the current context. First, Lee’s results are derived for a binary treatment while the $prcohort10_j$ measure as

derived from Bayes' rule and used throughout the paper is not binary. Second, two of the outcomes potentially subject to selection bias (unemployment status and occupational skill) have a discrete distribution. In regressions with covariates, it is therefore not clear which observations one would trim in the case where the p^{th} percentile falls between many identical outcome values.

To address the first problem (binary treatment), I use the binary form of $prcohort10_j$ derived using the modal age of entry assumption. Since the results are reasonably robust to this change, this should be sufficient to get a sense of the extent of the sample selection problem while allowing Lee's method to be applied directly. To address the discrete distribution of the outcome variable, I slightly modify the procedure suggested by Lee in his footnote 27 to take into account the presence of covariates so that observations with the same treatment status and outcome value are not identical. Specifically, I sort treated observations by the outcome variable and a randomly assigned number. I then drop observations until the sum of the sampling weights in the remaining observations equals the total sum of the sampling weights in the treatment group multiplied by the untrimmed proportion. In order to account for the random aspect of the trimming procedure as well as the difficulty in computing the standard errors of the bounds, I perform 500 bootstrap replications. The resulting distribution of each bound is then used to compute its mean and standard error.

Appendix Table A2 shows the bounds obtained by conducting this exercise for the two main variables potentially affected by sample selection bias: the occupational skill measure and log wages. For completeness I also consider the unemployment variable, although the scope for sample selection bias is low given that mean labor force participation is around 90 percent. The regressions differ from those in the main body of the paper since I use the binary version of the instrument and also adopt a linear probability model for unemployment in order to simplify the computations. Due to these slight modifications, the first row of the table provides a baseline by re-running the original regressions with these modifications. The next two rows of the table show the bootstrapped upper and lower bounds with their standard errors, while the final row shows the trimming percentile, p , in the original sample.

As noted in section b of this appendix, the use of a binary instrument tends to attenuate the basic estimates somewhat (first row), but they retain the same pattern and significance levels as before. For unemployment, the low potential for sample selection bias is evident in the fact that one needs to trim only at around the 6th percentile (column 1). The effect on unemployment for the full sample can therefore be bounded as between -10.2 and -4.6 percentage points, which still implies a highly statistically significant and negative effect even in the worst case scenario.

For occupational skill, the upper bound is 0.31 standard deviations (around double the original point estimate), but the lower bound is close to zero. The fact that the lower bound extends to zero effects is likely due to the particular distribution of occupational skill. Specifically, virtually no males in the sample who experienced a 9 year regime are employed in occupational skill level 3, while around 10% of those who experienced a 10 year regime are employed in this level. This is exactly the trimming proportion. In other words, the lower bound calculation assumes that *all* of the 10% of individuals in skill level 3 are marginal individuals who appear due to selection and drops these from the sample. Since the strongest impact of the longer regime for males was to increase the probability of working in skill level 3, dropping these individuals leads to near zero estimated treatment effects.

For log wages the bounds range from 10.8 to 24.6 percentage points (column 3), both of which are significant at the 5% level or better. These bounds are respectively 17% lower and 89% higher than the point estimate in the first row. Thus, even if the individuals subject to selection are drawn from the very top of the wage distribution, experiencing a 10 year regime still results in significantly higher wages for the male sample.

Intuitively it seems that the direction of any sample selection bias, if present, is likely to be downwards in this context so that the true effect is closer to the estimated upper bounds. First, consider again the unemployment variable, which may be affected by the endogeneity of labor force participation. Around 56% of individuals with 9 or 10 years of education who are not in the labor force report that their main activity is performing domestic chores in their own house for no gain (i.e. excluding forms of home production), and a decrease in this category is responsible for the bulk of the increase in

labor force participation associated with the change in regime. It is difficult to see why individuals who would have been performing domestic chores if subject to the shorter regime would have any advantage in the labor market that would lead to a higher probability of employment and hence cause an upwards bias in the unemployment estimates.

Next consider the occupational skill and log wage variables, which are available only for individuals who are currently employed and are therefore potentially affected by the endogeneity of employment status. For sample selection to cause upwards bias, marginal individuals who are only employed if subject to the 10 year regime would have to enter relatively high-skilled and highly-paid occupations. This could be the case if their reservation wage is high, so that they are only willing to accept employment when the offer wage exceeds their gains from some other kind of activity. However, the LFS questionnaire is extremely detailed about the types of activities that count as work, and is specifically designed to capture activities related to agriculture and the informal sector that are relevant in Botswana. This includes activities for barter or home use. It is therefore reasonable to assume that the unemployed are not engaged in some other gainful activity that the LFS fails to capture. Moreover, Botswana does not have a national unemployment insurance scheme that would yield some other benefit to being unemployed.

It therefore seems unlikely that large numbers of highly able individuals choose voluntary unemployment if subject to a 9 year regime but tend to enter high-skilled and highly-paid level occupations if they happen to experience a 10 year regime. Instead, the nature of the sample selection bias seems far more likely to involve selected individuals being lemons. In this case, the computed upper bounds imply that the true estimates for occupational skill and wages could be up to double the estimates presented earlier and suggest that the former should be viewed as conservative ones even leaving measurement error issues aside.

Table 1: Summary statistics

	ed>0 (1)	ed=7 (2)	ed=9 or 10 (3)	ed=12 (4)
Panel A: Demographic characteristics				
age	27.7 (6.0)	29.8 (6.2)	26.8 (5.5)	26.8 (5.5)
male	0.437 (0.496)	0.399 (0.490)	0.402 (0.490)	0.487 (0.500)
married	0.134 (0.341)	0.157 (0.364)	0.114 (0.318)	0.147 (0.354)
reside in Gaborone	0.161 (0.367)	0.126 (0.332)	0.144 (0.351)	0.254 (0.435)
years education	8.84 (2.59)	7.00 (0.00)	9.46 (0.50)	12.00 (0.00)
Panel B: Labor market outcomes				
LF participation	0.883 (0.321)	0.884 (0.320)	0.902 (0.297)	0.871 (0.335)
unemployed	0.351 (0.477)	0.373 (0.484)	0.363 (0.481)	0.286 (0.452)
occ. skill (std. dev.)	0.000 (1.000)	-0.459 (0.695)	-0.096 (0.763)	0.809 (1.144)
hourly wage (2006 Pula)	9.68 (9.27)	4.38 (2.74)	7.22 (5.74)	19.98 (14.59)
fraction of sample	1.00	0.22	0.37	0.25

Notes: Data are for the pooled sample using sampling weights. Data in panel A are available in all surveys, data in panel B are available in the LFS's only. For the hourly wage, observations above the 95th percentile or below the 5th percentile were dropped as outliers for each sample. Fractions of sample are relative to sample with non-zero educational attainment. Standard deviations are in parentheses. See text for details of variable definitions.

Table 2: Reduced form impact of structural regime on educational attainment

	(1)	(2)	(3)
Panel A: Educ>7			
Pr(cohort10)	0.322*** (0.024)	0.324*** (0.029)	0.364*** (0.068)
Pr(cohort10)*I _{>1973}			-0.095 (0.160)
Adj R ²	0.024	0.025	0.025
N	400606	400606	400606
Panel B: Educ=9 or 10			
Pr(cohort10)	0.620*** (0.035)	0.618*** (0.043)	0.528*** (0.064)
Pr(cohort10)*I _{>1973}			0.213* (0.110)
Adj R ²	0.298	0.330	0.332
N	225404	225404	225404
Survey year dummies	Y	Y	Y
Other controls	N	Y	Y
Linear yob trend	N	Y	Y

Notes: Each coefficient in the table is from a separate OLS regression of educational attainment on the probability of being in a 10 year regime (coefficient shown in the table) and controls as indicated. Column 3 includes the interaction between the probability of being in a 10 year regime and a post-1973 birth cohort dummy as indicated. “Other controls” are the number of individuals dropping out at junior secondary level and the student-teacher ratio in grade 8. All regressions are for the pooled sample using sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 3: Reduced form impact of structural regime on completing each year of education – sample with some secondary education (educ>7)

Dependent variable	(1)	(2)
I(educ>=9)	0.002 (0.002)	-0.003 (0.003)
I(educ>=10)	0.345*** (0.019)	0.341*** (0.024)
I(educ>=11)	-0.015* (0.009)	-0.01 (0.009)
I(educ>=12)	-0.01 (0.010)	-0.005 (0.010)
N	400606	400606
Survey year dummies	Y	Y
Other controls	N	Y
Linear yob trend	N	Y

Notes: Each coefficient in the table is from a separate OLS regression of a dummy for completing at least a certain level of education on the probability of being in a 10 year regime (coefficient shown in the table) and controls as indicated. “Other controls” are the number of individuals dropping out at junior secondary level and the student-teacher ratio in grade 8. All regressions are for the full pooled sample using sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 4: Reduced form effect of structural regime on labor force participation and unemployment - sample with 9 or 10 years of education

	(1)	(2)	(3)	(4)	(5)
LF participation	0.037*** (0.007)	0.040*** (0.007)	0.039*** (0.006)	0.050*** (0.009)	0.013 (0.009)
N	6587	6587	6587	4111	2476
Adj R ²	0.072	0.073	0.074	0.054	0.113
unemployed	-0.094*** (0.010)	-0.090*** (0.010)	-0.084*** (0.011)	-0.102*** (0.017)	-0.041* (0.021)
N	5870	5870	5870	3567	2303
Adj R ²	0.085	0.085	0.085	0.086	0.071
Cohort controls	N	N	Y	Y	Y
Age polynomial	cubic	quartic	cubic	cubic	cubic
Sample	full	full	full	females	males

Notes: Coefficients are from probit regressions of each outcome on the probability of being in a 10 year regime (coefficient shown in the table), a survey year dummy, dummies for male (in columns 1-3), married and Gaborone (capital city), a linear birth cohort trend and other controls as indicated. Coefficients are evaluated at the sample mean value of the covariates. "Cohort controls" are the number of individuals dropping out at junior secondary level, the student-teacher ratio in grade 8 and the growth rate of real GDP in the year of labor market entry. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 5: Reduced form effect of structural regime on occupational skill – male sample with 9 or 10 years of education

Dependent variable	(1)	(2)	(3)
occ. level 1	-0.046 (0.029)	-0.065** (0.026)	-0.040* (0.023)
occ. level 2	-0.055* (0.031)	-0.051* (0.029)	-0.064** (0.027)
occ. level 3	0.060*** (0.013)	0.062*** (0.014)	0.055*** (0.011)
occ. level 4	0.001 (0.002)	0 (0.000)	0 (0.000)
occ. skill (s.d)	0.203*** (0.049)	0.219*** (0.046)	0.191*** (0.044)
N	1560	1560	1560
Adj R ²	0.084	0.085	0.084
Cohort controls	N	N	Y
Age polynomial	cubic	quartic	cubic

Notes: Coefficients are from regressions of each outcome on the probability of being in a 10 year regime (coefficient shown in the table), a survey year dummy, dummies for male, married and Gaborone (capital city), a linear birth cohort trend and other controls as indicated. For the occupational skill level dummies (levels 1-4) the regressions are probits and the coefficients are evaluated at the sample mean value of the covariates. For occupational skill (in standard deviations), the regressions are OLS. “Cohort controls” are the number of individuals dropping out at junior secondary level, the student-teacher ratio in grade 8 and the growth rate of real GDP in the year of labor market entry. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 6: Reduced form effect of structural regime on log wages – male sample with 9 or 10 years of education

	(1)	(2)	(3)
lnwage	0.134*** (0.052)	0.134*** (0.052)	0.158*** (0.058)
N	1151	1151	1151
Adj R ²	0.322	0.321	0.328
Cohort controls	N	N	Y
Age polynomial	cubic	quartic	cubic

Notes: Coefficients are from OLS regressions of log wages on the probability of being in a 10 year regime (coefficient shown in the table), a survey year dummy, dummies for male, married and Gaborone (capital city), a linear birth cohort trend and other controls as indicated. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers. “Cohort controls” are the number of individuals dropping out at junior secondary level, the student-teacher ratio in grade 8 and the growth rate of real GDP in the year of labor market entry. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 7: Contribution of occupational choice to reduced form effect of structural regime on log wages – male sample with 9 or 10 years of education and reporting occupational skill level

	(1)	(2)	(3)
lnwage	0.143** (0.059)	0.104** (0.053)	0.085 (0.054)
N	1121	1121	1121
Adj R ²	0.323	0.354	0.362
occ. skill (s.d.)	N	Y	N
occ. skill dummies	N	N	Y

Notes: Coefficients are from OLS regressions of log wage on the probability of being in a 10 year regime (coefficient shown in the table), a survey year dummy, dummies for male, married and Gaborone (capital city), a linear birth cohort trend, the number of individuals dropping out at junior secondary level, the student-teacher ratio in grade 8 and the growth rate of real GDP in the year of labor market entry. This is the same specification as column 3 of Table 6. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table 8: OLS and IV estimates of returns to the 10th year of education – male sample

	(1)	(2)	(3)
	Panel A: OLS		
ln(wage)	0.175*** (0.058)	0.175*** (0.058)	0.231*** (0.062)
	Panel B: IV		
ln(wage)	0.192** (0.075)	0.187*** (0.072)	0.263*** (0.098)
Cohort controls	N	N	Y
Age polynomial	cubic	quartic	cubic
Partial R ²	0.316	0.322	0.212
Partial F	223.3	248.7	211.3
N	1151	1151	1151

Notes: Coefficients in panel A are from OLS regressions of log wages on years of education (coefficient shown in the table), a survey year dummy, dummies for male, married and Gaborone (capital city), a linear birth cohort trend and other controls as indicated. “Cohort controls” are the number of individuals dropping out at junior secondary level, the student-teacher ratio in grade 8 and the growth rate of real GDP in the year of labor market entry. Coefficients in panel B are from analogous 2SLS regressions using the probability of being in a 10 year regime as an instrument for years of education. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table A1: Reduced form estimates using different constructions of the instrument – sample with 9 or 10 years of education

	Bayes' rule			Modal age		
	No repeat	Repeat Pr=50%	Repeat Pr=100%	No repeat	Repeat Pr=50%	Repeat Pr=100%
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Education						
education	0.618*** (0.043)	0.639*** (0.037)	0.634*** (0.036)	0.541*** (0.058)	0.592*** (0.057)	0.539*** (0.072)
Panel B: Labor market						
LF participation	0.039*** (0.006)	0.039*** (0.007)	0.039*** (0.008)	0.034*** (0.006)	0.035*** (0.007)	0.031*** (0.008)
unemployed	-0.084*** (0.011)	-0.092*** (0.010)	-0.098*** (0.010)	-0.071*** (0.012)	-0.085*** (0.009)	-0.090*** (0.008)
occ. skill	0.191*** (0.044)	0.206*** (0.044)	0.217*** (0.043)	0.158*** (0.048)	0.187*** (0.042)	0.191*** (0.036)
ln(wage)	0.158*** (0.058)	0.174*** (0.059)	0.188*** (0.061)	0.130*** (0.044)	0.162*** (0.041)	0.173*** (0.039)

Notes: Coefficients are from regressions of each outcome on the probability of being in a 10 year regime (coefficient shown in table) and a set of controls. “Bayes’ rule” refers to the calculation of the probabilities as used in the paper. “Modal age” refers to the assumption that students enter school at age 7. For Panel A, the controls are as in column 3 of Table 2. For Panel B, the controls are as in column 3 of Tables 4-6. For binary outcomes the regressions are probits and the coefficients are evaluated at the sample mean. For non-binary outcomes the regressions are OLS. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers for the log wage regressions. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table A2: Bounds for sample selection bias – sample with 9 or 10 years of education

	unemployed (1)	occ. skill (s.d) (2)	lnwage (3)
estimate	-0.065*** (0.011)	0.158*** (0.048)	0.130*** (0.044)
lower bound	-0.046*** (0.016)	0.026 (0.053)	0.108** (0.050)
upper bound	-0.102*** (0.017)	0.306*** (0.070)	0.246*** (0.065)
p th percentile	5.5	9.9	12.1
sample	full	male	male

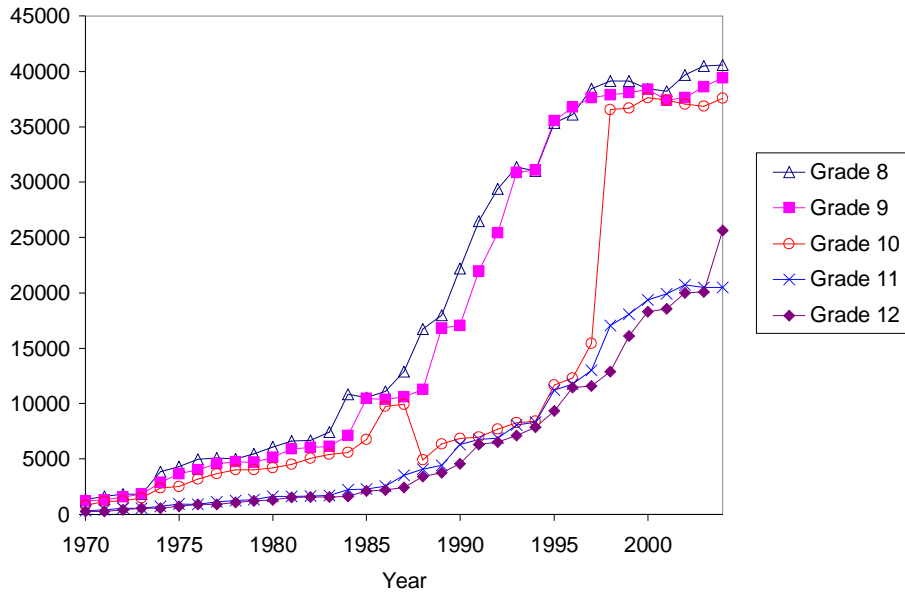
Notes: Coefficients in first row are from an OLS regression of each outcome on the probability of being in a 10 year regime using the modal age of entry assumption and a set of controls that are the same as those in column 3 of Tables 4-6. Lower bounds and upper bounds (Lee, 2008) were computed using 500 bootstrap replications. The trimming percentile in the original sample is denoted by “pth percentile” and is calculated as described in the text. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers for the log wage regressions. All regressions use sampling weights and those in the first row have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Table A3: Reduced form estimates by gender – sample with 9 or 10 years of education

	(1) Males	(2) Females
Panel A		
occ. level 1	-0.040* (0.023)	-0.105*** (0.020)
occ. level 2	-0.064** (0.027)	0.037 (0.035)
occ. level 3	0.055*** (0.011)	0.044 (0.030)
occ. level 4	0 (0.000)	0 (0.000)
occ. skill (s.d)	0.191*** (0.044)	0.262*** (0.046)
Panel B		
lnwage	0.158*** (0.058)	0.219*** (0.051)

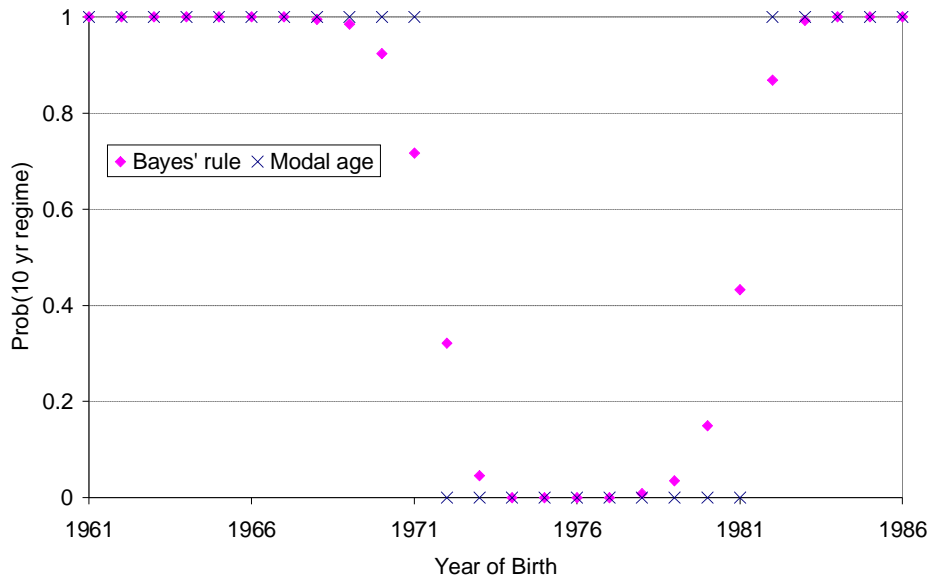
Notes: Coefficients are from regressions of each outcome on the probability of being in a 10 year regime (coefficient shown in the table) and a set of controls. For Panel A the controls are as in column 3 of Table 5 and for Panel B the controls are as in column 3 of Table 6. For binary outcomes the regressions are probits and the coefficients are evaluated at the sample mean value of the covariates. For non-binary outcomes, the regressions are OLS. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers for the log wage regressions. All regressions use sampling weights and have robust standard errors clustered at the birth cohort level. Standard errors are in parentheses. *= significant at 10% **= significant at 5%, ***= significant at 1% levels.

Figure 1: Secondary school enrollment by grade



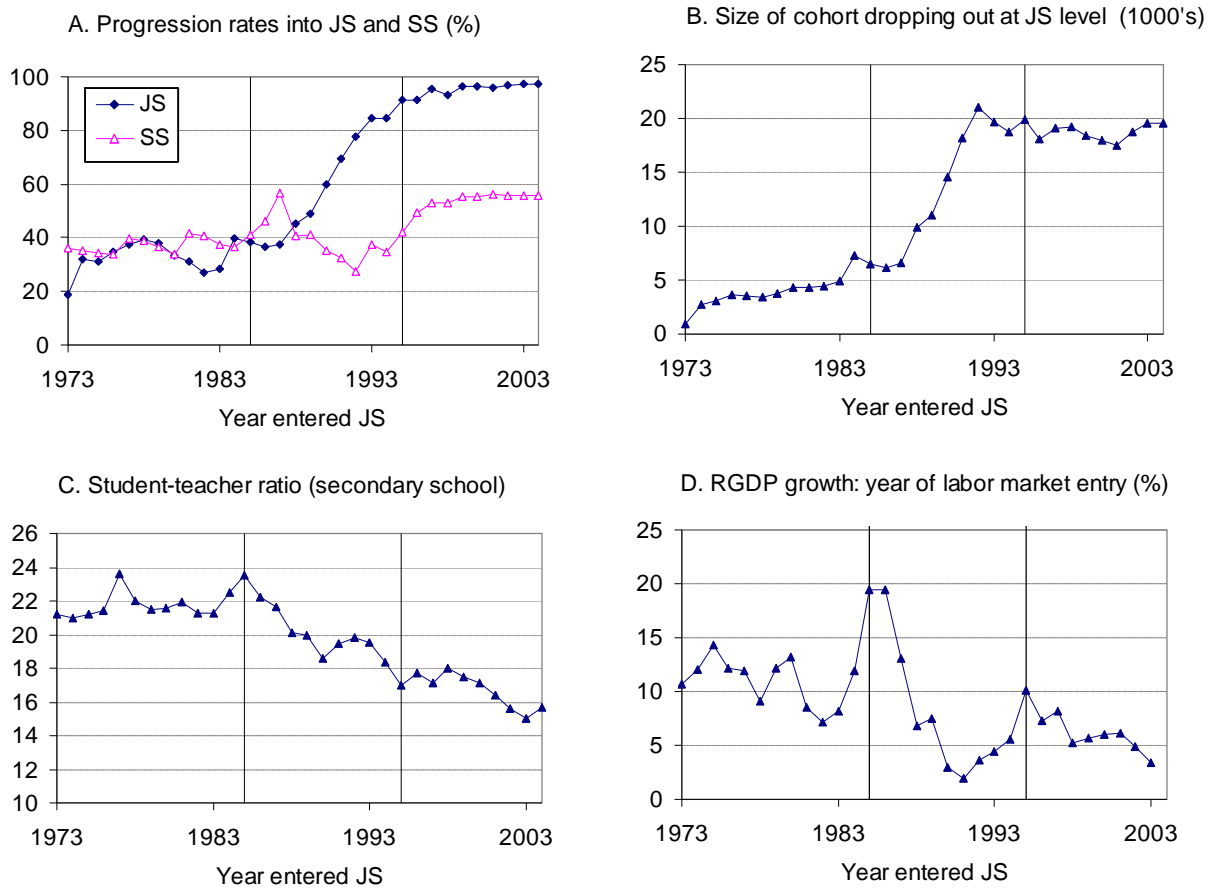
Notes: Data are from annual *Education Statistics Reports* and include government schools, community-run “unaided” schools and private schools.

Figure 2: Calculated probability of experiencing a 10 year regime using Bayes’ rule or modal age of entry assumption.



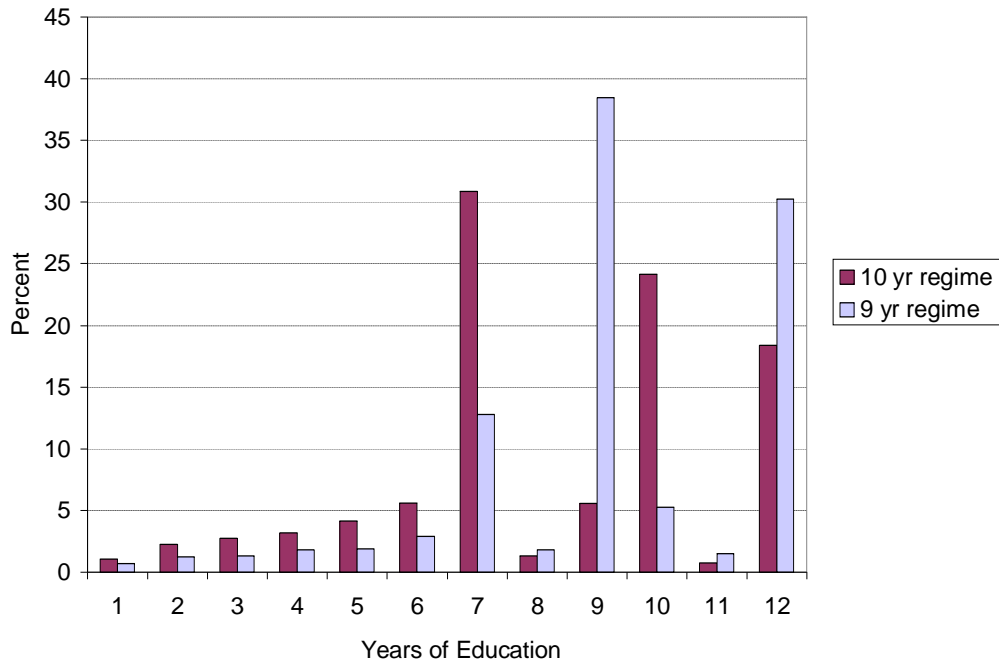
Notes: Bayes’ rule calculation uses birth cohort and administrative information on the distribution of age of entry and cohort size as described in the text (section 3). Modal age calculation assumes school entry at age 7.

Figure 3: Key contemporaneous changes by educational cohort.



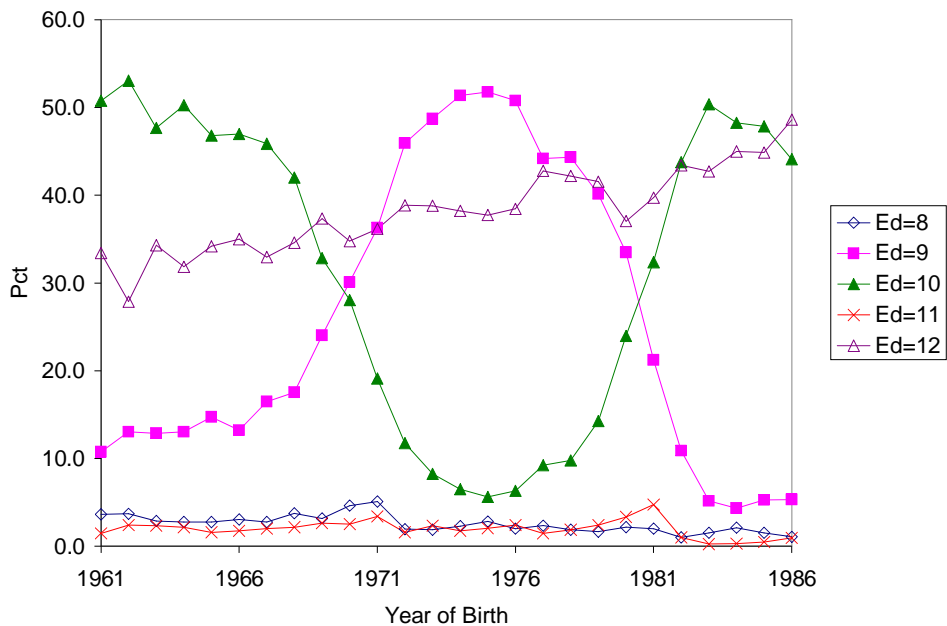
Notes: Data for panel A-panel C are from annual *Education Statistics Reports* and include government schools, community-run “unaided” schools and private schools. Data from panel D are from the World Development Indicators (World Bank, 2008). JS denotes Junior Secondary and SS denotes Senior Secondary. Panel A: progression rate is enrollment in grade of entry as a fraction of enrollment in the preceding grade in the previous year. Panel B: size of cohort dropping out at JS level is calculated as enrollment in grade of entry to SS as a fraction of grade 8 enrollment (first year of JS) for each educational cohort. Panel C: student-teacher ratio is calculated in year cohort enters grade 8. Panel D: real GDP growth rate is based on GDP figures in constant year 2000 dollars with the year of labor market entry calculated assuming completion of JS. In each panel, the 9 year regime affects cohorts between the two vertical lines.

Figure 4: Histogram of educational attainment conditional on educ>0



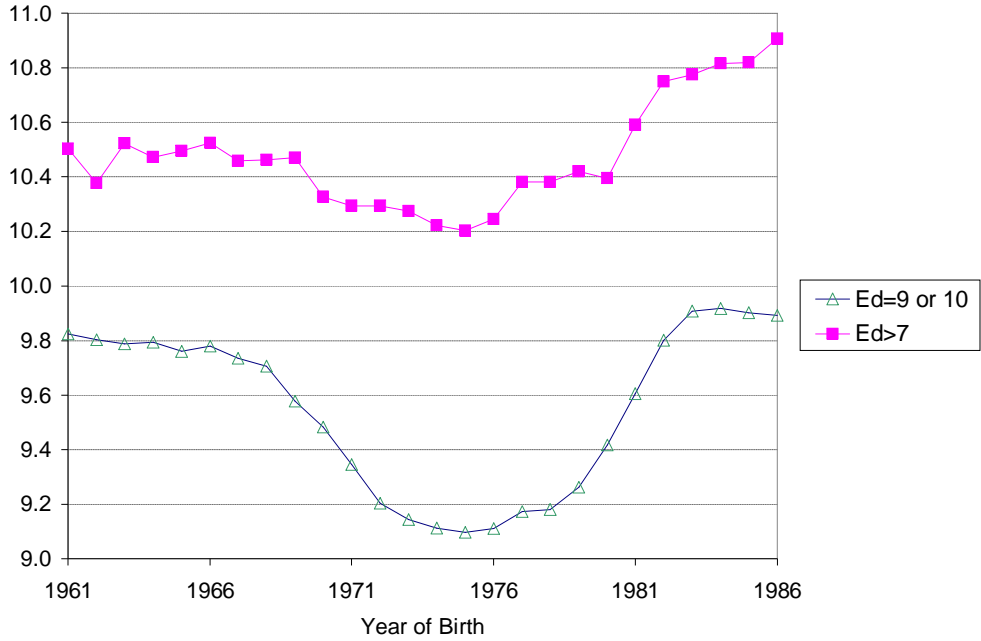
Notes: Data are from full pooled sample using sampling weights. Regime determined by application of Bayes' rule: those for whom the probability of a 10 year regime is not zero or unity are omitted.

Figure 5: Distribution of educational attainment conditional on educ>7



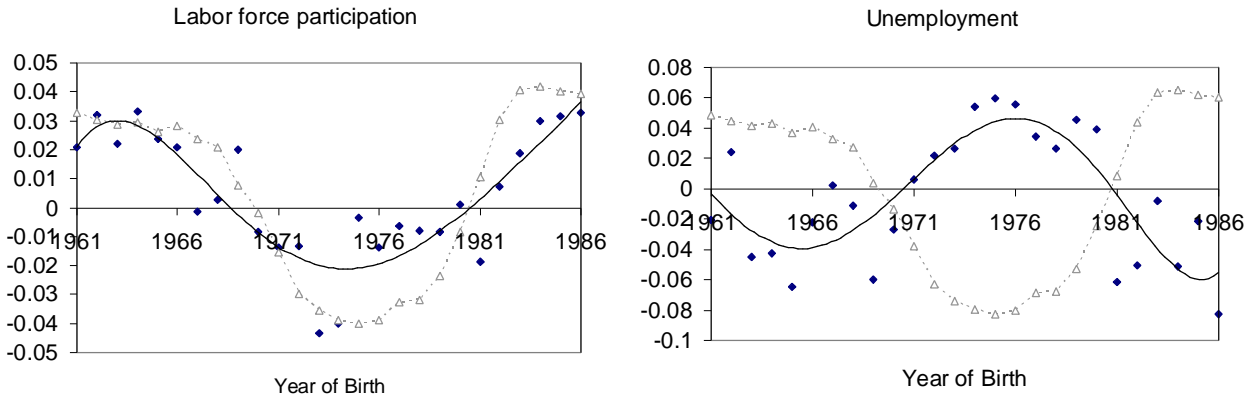
Notes: Data are from full pooled sample using sampling weights.

Figure 6: Mean educational attainment



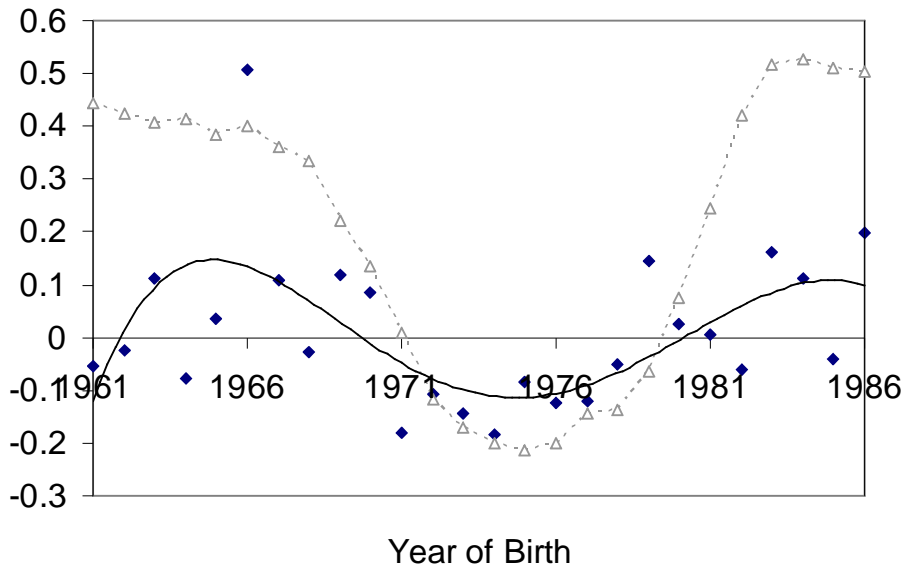
Notes: Data are from pooled sample using sampling weights.

Figure 7: Mean residuals for labor force participation and unemployment by birth cohort - sample with 9 or 10 years of education.



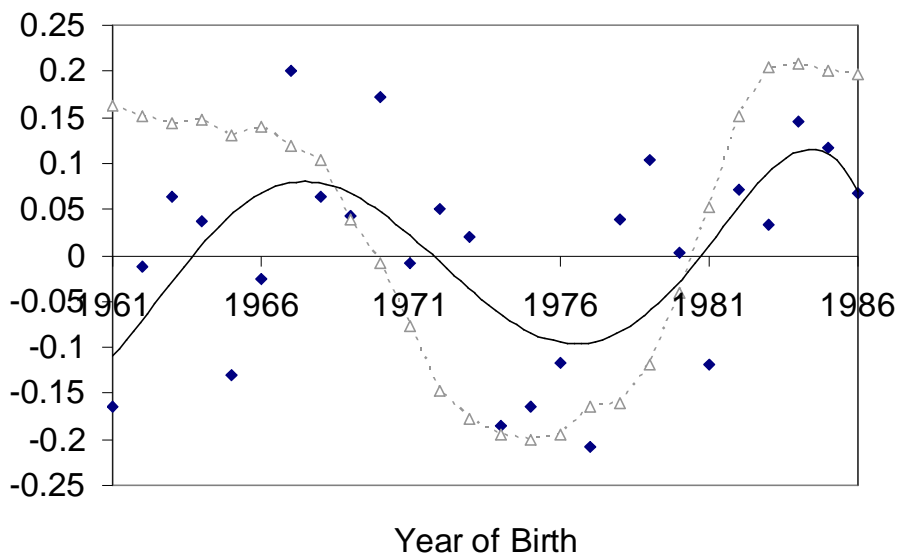
Notes: Solid markers plot mean residuals after a probit on a cubic in age and a survey year dummy for the sample with 9 or 10 years of education. Sampling weights are used. A fifth order polynomial (solid line) is fitted to aid the visual impression. Hollow markers (joined with dashed line) show the pattern of mean educational attainment for reference.

Figure 8: Mean residuals for the normalized occupational skill measure (in standard deviations) by birth cohort - males with 9 or 10 years of education.



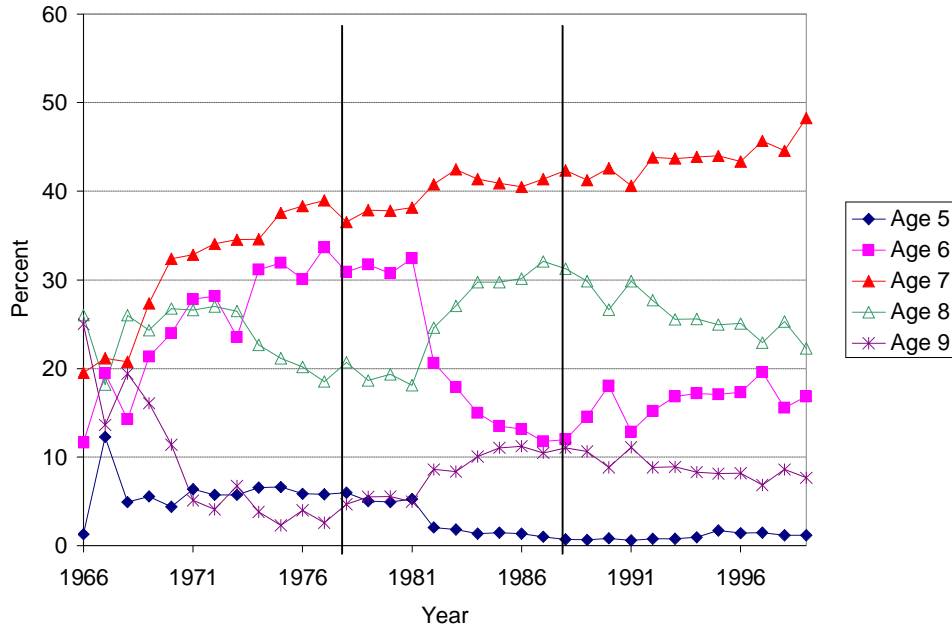
Notes: Solid markers plot mean residuals after a regression on a cubic in age and a survey year dummy for the sample with 9 or 10 years of education. Sampling weights are used. A fifth order polynomial (solid line) is fitted to aid the visual impression. Hollow markers (joined with dashed line) show the pattern of mean educational attainment for reference.

Figure 9: Mean residuals for log wages by birth cohort - males with 9 or 10 years of education.



Notes: Solid markers plot mean residuals after a regression on a cubic in age and a survey year dummy for the sample with 9 or 10 years of education. Observations with log wages above the 95th percentile or below the 5th percentile were dropped as outliers. Sampling weights are used. A fifth order polynomial (solid line) is fitted to aid the visual impression. Hollow markers (joined with dashed line) show the pattern of mean educational attainment for reference.

Figure A1: Distribution of age in year of school entry



Notes: Data are from annual *Education Statistics Reports*. Vertical lines demarcate changes in structural regime: the 9 year regime affects those entering in years between the two lines.