Initially, I examined whether the running variable was correlated with predetermined covariates, specifically, the language of interview and the respondent's relationship to the household head. The language variable includes categories such as Zulu, Xhosa, Ndebele, Swati, Sesotho, Setswana, Tsonga, Venda, English, Afrikaans, and Sepedi. The relationship to household head variable includes head, spouse, child, sibling, parent, grandchild, grandparent, other relative, and non-relative. Upon analysis, I found that some language and relationship categories were significantly associated with the running variable, suggesting potential imbalance. As a result, I included these covariates in the relevant models to improve the validity of the RD estimates.

4.4 Results

This section presents the main findings on the effects of education on labour market earnings using school entry age as an instrument for education in a regression discontinuity design. Results are shown for the overall sample, and separately by gender, over the period 2014–2018. The section proceeds with descriptive statistics, manipulation checks, RD estimates for educational attainment and log earnings, followed by robustness and sensitivity tests.

The GHS 2014-2018 has repeated observations throughout the period, however, not all individuals were interviewed in each of the years, making it an unbalanced panel. To avoid the bias, which may arise due to the unbalanced panel nature of the data, I provide RD estimates for each survey year. For comparison purposes, I also show results for the pooled sample. Two estimation strategies are undertaken, the standard fuzzy RD, which may be affected by manipulation around the cut-off, and the donut RD, which excludes observations near the cut-off to correct for potential manipulation.

4.4.1 Descriptive statistics

In this section I present and discuss the descriptive statistics for both men and women (see Table 4.1). The total sample consists of 69,673 individuals of which 37,352 (53.6%) are men and 32,321 (46,4%) are women. There are more men than women in the sample because there is a higher proportion of men who are employed compared to women. However, it should be noted that the actual number of observations used in the econometric analysis will be limited by the bandwidths. Overall, the treated group consist of 35,614 (51.1%) individuals compared to 34,059 (48.9%) workers in the control group. A comparison is made between the control and treated groups.

Table 4.1: Summary statistics of key variables for all individuals, women and men

	All individuals			Women			Men		
	Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
	Mean	Mean	b	Mean	Mean	b	Mean	Mean	b
Age (years)	39.02	39.46	-0.44***	39.88	40.34	-0.46***	38.25	38.71	-0.46***
Education (years)	10.38	10.08	0.30***	10.59	10.29	0.30***	10.20	9.91	0.29***
Real Wage (Rands)	6079.32	5833.01	246.31***	5649.81	5467.62	182.19**	6461.40	6140.92	320.47***
Female	0.47	0.46	0.01***	1.00	1.00	0.00	0.00	0.00	0.00
Legally married	0.31	0.32	-0.01**	0.27	0.28	-0.01*	0.34	0.34	-0.00
Metropolitan	0.39	0.38	0.01***	0.39	0.37	0.01**	0.40	0.38	0.02***
Household size	3.76	3.80	-0.03*	4.18	4.26	-0.08***	3.39	3.41	-0.02
Western Cape	0.06	0.06	0.00	0.06	0.05	0.01***	0.06	0.06	-0.00
Eastern Cape	0.23	0.23	0.00	0.26	0.25	0.01	0.21	0.21	-0.00
Northern Cape	0.03	0.03	-0.00	0.03	0.03	0.00	0.03	0.04	-0.00*
Free State	0.08	0.07	0.00**	0.07	0.07	0.00	0.08	0.07	0.01***
KwaZulu-Natal	0.16	0.16	0.01**	0.17	0.17	0.00	0.16	0.15	0.01**
North West	0.07	0.08	-0.00*	0.07	0.07	0.00	0.08	0.08	-0.01**
Gauteng	0.28	0.27	0.01**	0.26	0.26	0.01	0.29	0.28	0.01**
Mpumalanga	0.11	0.11	-0.00	0.10	0.11	-0.01**	0.11	0.11	0.00
Limpopo	0.10	0.11	-0.02***	0.10	0.12	-0.02***	0.09	0.11	-0.02***
Observations	34059	35614	69673	16034	16287	32321	18025	19327	37352

Note: Difference represents the difference in means between the control and treated groups. b is the coefficient of the difference in means between the control and treated groups. ***, **, and * indicate statistical significance at the 1, 5, and 10% critical levels.

The treated group is slightly older compared to the control group with a statistically significant difference of 0.44 years between them. However, the latter is more educated with an additional 0.3 years of schooling compared to the former. Hence, the control group has higher earnings with an additional R246.31, on average per month compared to the treated group. Additionally, the control group has a higher proportion of workers living in metropolitan areas. The earnings of women and men mirrors that of the overall sample with the control group earning higher on average compared to their treated counterparts. As with the overall sample, this difference is statistically significant seems to be driven by differences in years of education as one would expect.

Table 4.1 indicates also that there are gender differences in wages with women earning less than men, on average, despite having more years of education, on average. The monthly average earnings for untreated women is only 87.4% of untreated men's average monthly earnings whilst the monthly average earnings of treated women is just 89.0% of their treated male counterparts. One possible reason for this difference in wages could be the over-representation of women in low-paying jobs due to a historical patriarchal society (Espi et al., 2019; Mosomi, 2019*b*). It is observable also that there are differences in the proportion of treated and untreated women and men across the 9 provinces.

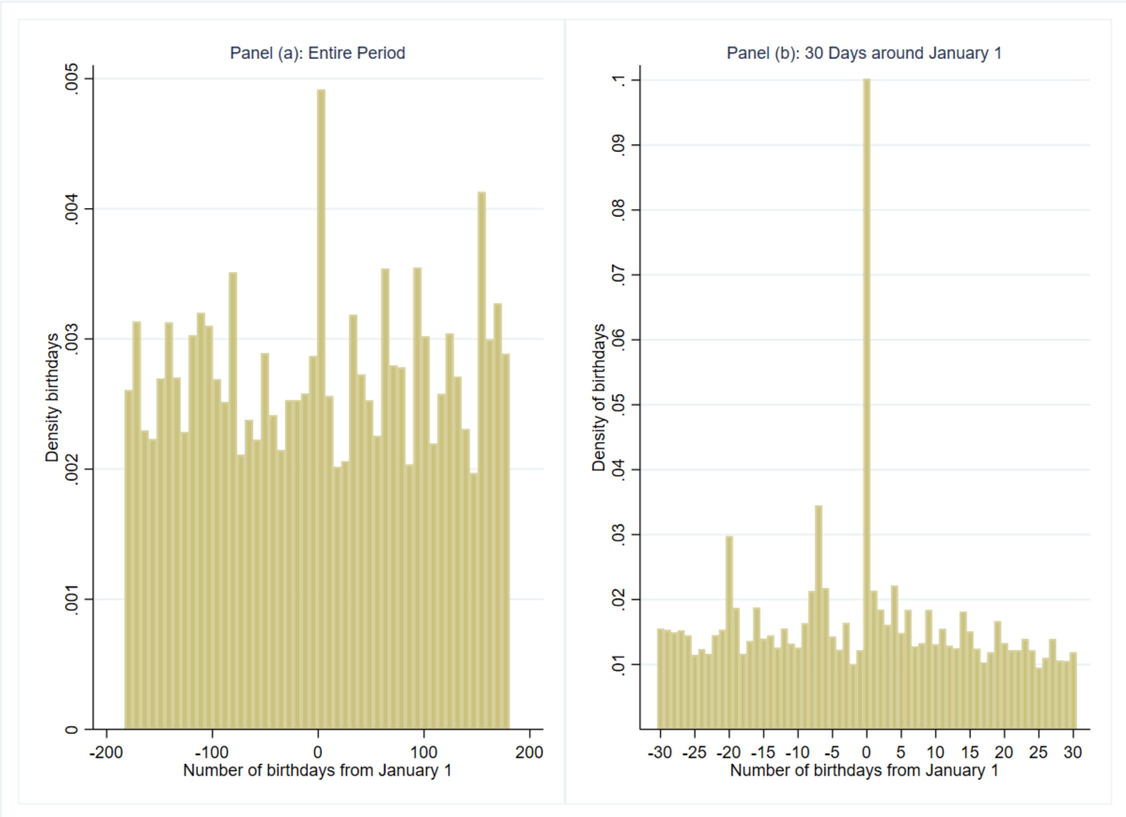
4.4.2 Evidence of manipulation

As stated earlier, the RD design is based on the assumption that birthdays, which is the running variable, cannot be manipulated by participants. However, this analysis shows evidence of manipulation. To illustrate this strategic behaviour, I present two histograms displaying the distribution of birth dates. One histogram shows the distribution of birth dates for the overall (or full) sample, while the other focuses on the period between 1 December and 30 January, where manipulation is suspected to have occurred. Figure 4.1 shows the two histograms side-by-side. Panel (a) of the figure depicts the histogram of birth dates for the entire sample while panel (b) focuses on the suspected period covering 30 days before and after the actual cut-off date (1 January).

Panel (a) shows heaping of birth date at the cut-off. However, to show a more clear picture of the problem I present a histogram depicting birthdays within 30 days around 1 January on either side of the threshold. Panel (b) indicates a clear evidence of heaping on January 1, which implies that some births were possibly misreported or registered to fall just at the cut-off, likely to meet the school entry age requirements. This behaviour is non-random and violates the RD assumption of no precise control over the running variable near the cut-off. I conduct the McCrary manipulation test to formally investigate the issue.

This test involves examining the distribution of the running variable (birthdays) around the threshold to ensure that there is no manipulation or sorting just above or below the cut-off. The McCrary density test involves testing for a significant jump in the density of the running variable at the threshold (McCrary, 2008). A smooth density function indicates that individuals did not manipulate their position relative to the threshold. The test is implemented using the local

Figure 4.1: Birth date heaping around January 1

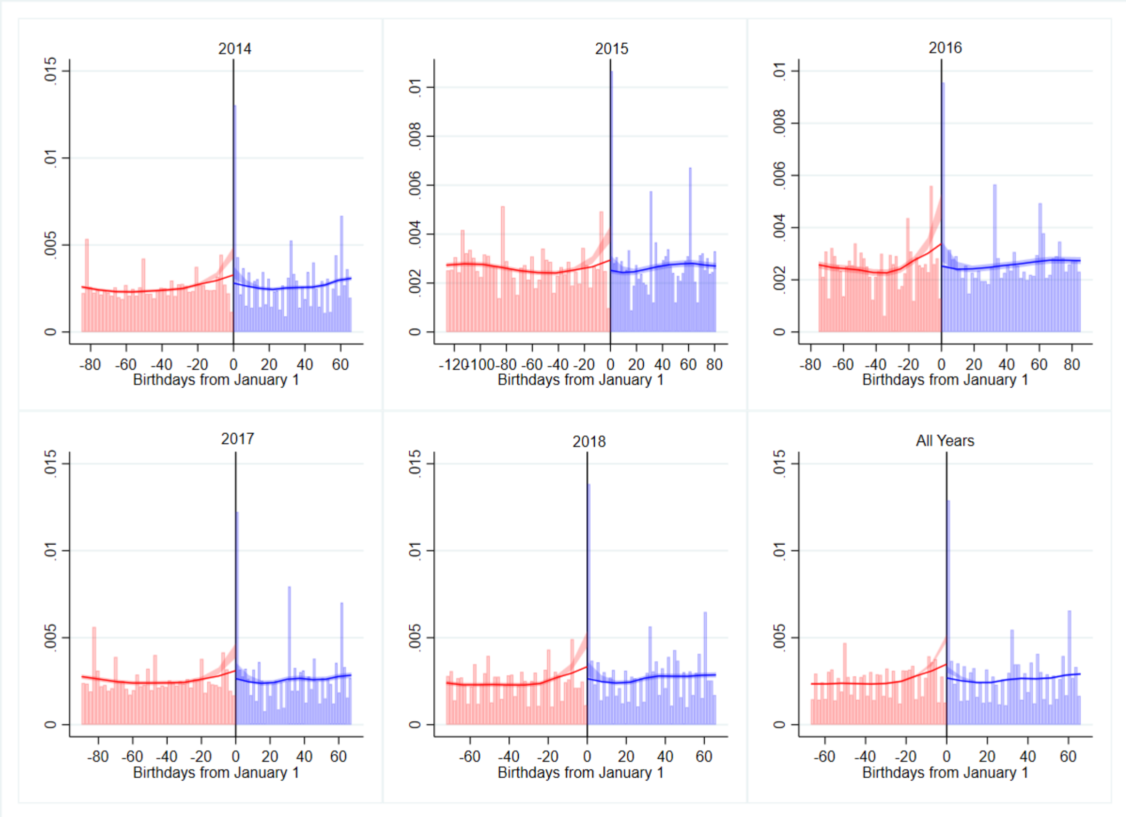


Source: Author's calculations based on the 2014-2018 GHS data.

polynomial density estimators proposed in Cattaneo, Jansson and Ma (2020). It is designed to test the continuity assumption around the cut-off point. It offers a clear and easy way to determine empirically if units near the cut-off point can change their assignment to treatment.

As noted before, units are assigned to the control group if born on or between 1 July and 31 December (or $x < x_0$) and to the treatment group if born between 1 January and 30 June (or $x \geq x_0$). The main aim is to test whether the density of birthdays is continuous at the cut-off point (1 January) using the control and treatment groups. Thus, the null hypothesis is that the density of birthdays is continuous around the cut-off point (no manipulation) against the alternative hypothesis that the density is discontinuous (manipulation) (Cattaneo, Jansson and Ma, 2020). The null hypothesis is rejected if the p-value related to the test is less than a chosen level of significance. The results are depicted in Figure 4.2. Figure 4.2 shows a big jump just at January 1 for each of the survey years, suggesting that birthdays is not continuous around the

Figure 4.2: Manipulation test for birthdays: 2014-2018



Note: Source: Author's calculations based on data from the GHS 2014-2018.

Table 4.2: McCrary manipulation test

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Full sample						
Birthdays	1.476	1.502	-0.686	0.701	1.549	1.106
	[0.140]	[0.133]	[0.493]	[0.484]	[0.121]	[0.269]
Female sample						
Birthdays	-1.384	-1.823*	-1.516	-1.581	0.121	-1.984**
	[0.166]	[0.068]	[0.130]	[0.114]	[0.903]	[0.047]
Male sample						
Birthdays	1.488	-2.307**	-1.062	1.975**	0.971	3.6955
	[0.137]	[0.021]	[0.288]	[0.048]	[0.332]	[0.000]

Note: The coefficients for birthdays are robust test statistic for the McCrarry density test. Robust p-values in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

cut-off. However, a formal density test suggests otherwise. The test results are depicted in Table 4.2 for each year under consideration as well as for all years combined. For the overall sample, the results indicate that the null hypothesis of no manipulation is not rejected for each of the years as well as for the pooled sample, suggesting that manipulation is not a problem. Similarly, for the male and female sub-samples, the p-value of the test in each year is greater than the conventional significance levels except for 2015. The fact that the p-value is greater than the level of significance suggests that the null hypothesis of continuity at the cut-off point cannot be rejected, indicating that birthdays is not manipulated.

Although the formal density test is not statistically significant for the overall sample in each year and for the sub-samples in most years, the visible heaping in birth dates at the cut-off is consistent with strategic behaviour by parents. Although the McCrary test does not reject the null hypothesis of continuity in density, its power is limited with discrete running variables (Barreca et al., 2011). Given the evidence of potential manipulation around the cut-off, I adopt a conservative approach by implementing a donut RD design. Following best practices in the literature (Barreca et al., 2011; Fan et al., 2018), I exclude observations with a date of birth exactly on the cut-off (January 1). To assess the robustness of the results, I extend the exclusion window incrementally to ± 1 and ± 2 days around the cut-off. For transparency, I report both the baseline RD estimates including all observations and the donut RD estimates that account for potential manipulation.

4.4.3 Impact on Education

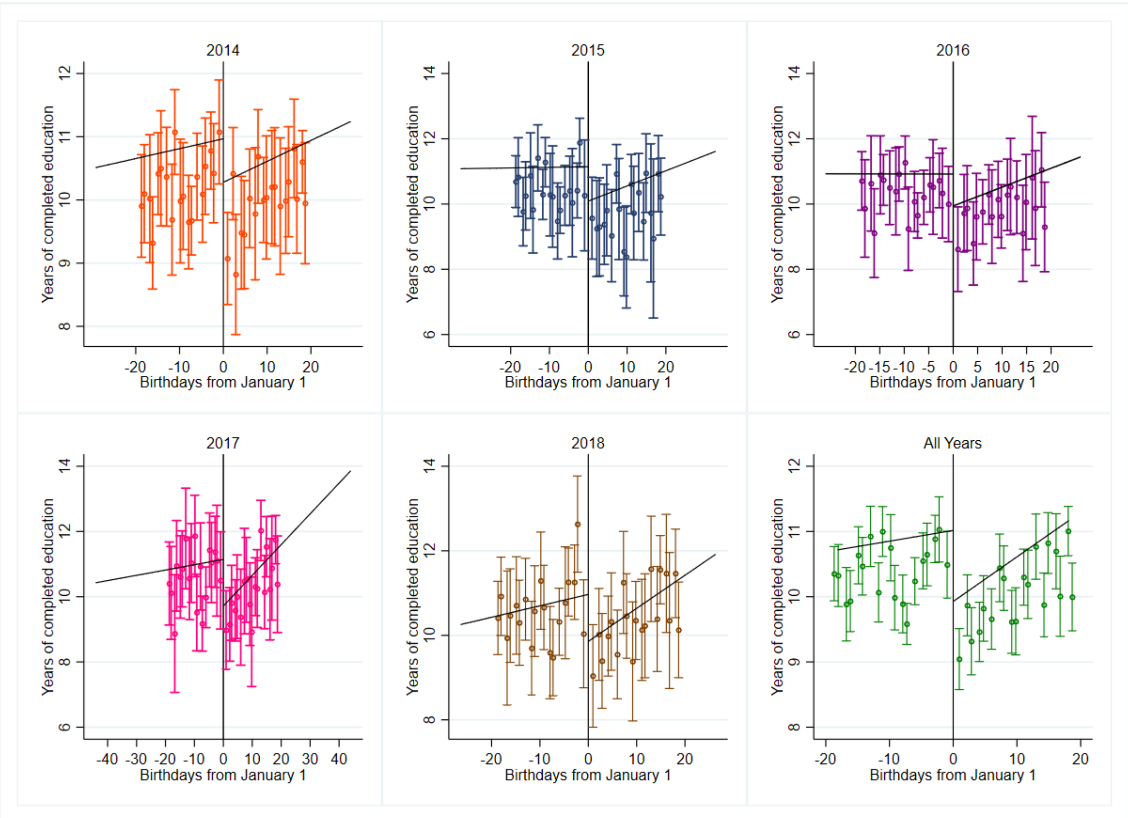
As outlined in the estimation strategy, the fuzzy RD design employs school entry age as an instrumental variable for years of education. School entry age is proxied by the running variable, the number of birthdays from the school entry cut-off (January 1). For the fuzzy RD approach to be valid, the first stage, the effect of school entry age on years of education, must be statistically significant to justify estimating the second stage, which captures the impact of years of education on earnings. Accordingly, I estimate the relationship between school entry age (SEA) and years of education to assess the strength of the instrument. Since province, marital status, age group, and sex exhibit discontinuities at the cut-off (see Appendix C), I include these variables as controls in the analysis. Additionally, where appropriate, I control for the language of interview and the respondent's relationship to the household head. I first present the overall sample estimates, followed by a gender-disaggregated discussion.

Overall Sample

The results for the overall sample is presented in Figure 4.3 and Table 4.3. The vertical axis of the graph measures completed years of education, while the horizontal axis shows the number of birthdays away from January 1, the school entry cut-off. The vertical line at zero divides the sample into two groups: the control group (to the left) and the treatment group (to the right, including zero).

The solid curves on either side of the cutoff represent regression lines for the control and treatment groups, estimated using Stata's rdplot command. The bars depict 95% confidence intervals. The graph shows that early school entrants (younger cohort) tend to have fewer years of completed education than later entrants (older cohort), indicating a negative relationship between SEA and educational attainment. Notably, the pooled effect across all years is larger than the individual yearly estimates. The corresponding numerical estimates are provided in Table 4.3. Panel (a)

Figure 4.3: Impact of school entry age on years of completed education for the overall sample



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

presents fuzzy RD estimates from the model potentially affected by birth date manipulation, while Panel (b) reports estimates after correcting for this issue by excluding individuals born on January 1. Both models control for sex, marital status, province, age group, language of interview (Afrikaans, Sesotho, Siswati, Xhosa), and relationship to the household head, covariates selected based on prior literature and data availability.

The corrected estimates in Panel (b) consistently yield smaller effects of SEA on years of education compared to the uncorrected results, both for individual years and in the pooled sample. For example, in the corrected model, the coefficient on SEA ranges from –0.771 in 2017 to –1.066 in 2018, whereas the uncorrected estimates range from –1.121 in 2016 to –1.438 in 2018. This discrepancy highlights the bias introduced by birth date manipulation around the cutoff.

Across all years and model specifications, the SEA coefficient is negative and statistically significant at the 1% level. The corrected estimates suggest that starting school one year earlier reduces completed years of education by approximately 9.3 to 12.8 months, depending on the year. This result aligns with international evidence (Kawaguchi, 2011; Li et al., 2022; Valdés and Requena, 2024; Marcenaro-Gutierrez and Lopez-Agudo, 2021; Solli, 2017). For instance, Kawaguchi (2011) find that in Japan, males who started school later accumulated 13.20 years of education, compared to 13.03 years among younger entrants. Similarly, Valdés and Requena (2024) report that in Spain, later school starters gained, on average, 0.2 additional years of education.

Older entrants likely benefit from greater maturity and skill development before starting school (Valdés and Requena, 2024). Conversely, although younger children start earlier, they tend to face higher repetition rates, leading to lower overall attainment (Matta et al., 2016). In South Africa, van der Berg et al. (2025) show that younger school entrants experience higher repetition rates, for example, in the Eastern Cape, repetition rates among younger boys reach 31%, compared to 19% for girls in the same age group. These high repetition rates are often linked to poor academic performance, increasing the risk of dropout and reducing educational achievement. The fuzzy RD estimates corrected for manipulation (Panel (b)) consistently show smaller effects of school entry age (SEA) on years of completed education compared to the uncorrected estimates, across all

Table 4.3: Impact on years education and logarithm of monthly earnings for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
	Panel (a): Fuzzy RD with potential manipulation					
First stage						
School entry age	-1.418***	-1.318***	-1.121***	-1.333***	-1.438***	-1.530***
	(0.266)	(0.302)	(0.260)	(0.257)	(0.304)	(0.153)
Reduced Form						
School entry age	-0.296***	-0.226**	-0.193**	-0.270***	-0.066	-0.242***
	(0.107)	(0.097)	(0.094)	(0.078)	(0.089)	(0.053)
Second stage						
Education	0.209***	0.171**	0.172**	0.203***	0.046	0.158***
	(0.070)	(0.067)	(0.079)	(0.056)	(0.060)	(0.031)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	22.229	29.524	33.947	45.208	26.867	18.586
Bandwidth Bias (b)	62.838	72.980	89.571	112.887	66.239	64.321
	Panel (b): Fuzzy RD corrected for manipulation (donut RD)					
First stage						
School entry age	-0.817***	-0.984***	-1.039***	-0.771***	-1.066***	-1.221***
	(0.267)	(0.345)	(0.343)	(0.285)	(0.351)	(0.186)
Reduced Form						
School entry age	-0.246**	-0.105	-0.217*	-0.156*	-0.070	-0.225***
	(0.101)	(0.109)	(0.126)	(0.086)	(0.106)	(0.064)
Second stage						
Education	0.302**	0.106	0.209*	0.203*	0.065	0.184***
	(0.133)	(0.102)	(0.114)	(0.104)	(0.098)	(0.049)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	33.679	32.055	27.804	45.264	26.352	18.575
Bandwidth Bias (b)	84.208	85.749	80.704	112.318	66.225	63.253

Note: The dependent variable in the first-stage equation is years of schooling while the dependent variable is the reduced form and second stage equations is the logarithm of earnings. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. The bandwidth size is measured in birthdays away from 1 January, on either side of the cut-off point. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*). Covariates include province of residency, sex and marital status. The symbols ✗ denotes no, and ✓ stands for yes.

years and in the pooled results. In the first-stage equation, the coefficient of SEA in the corrected model ranges from –0.771 in 2017 to –1.066 in 2018, while the uncorrected estimates range from –1.121 in 2016 to –1.438 in 2018. This notable difference underscores the impact of birth date manipulation around the cut-off on estimated treatment effects.

The coefficient of SEA is negative and statistically significant at the 1% level of significance in the two specifications and throughout the survey years. Therefore, the estimates suggest that the school entry age reduced the years of completed education by between 9.3 and 12.8 months between 2014 and 2018. This finding aligns with the literature (Kawaguchi, 2011; Li et al., 2022; Valdés and Requena, 2024; Marcenaro-Gutierrez and Lopez-Agudo, 2021; Solli, 2017). For example, the findings of Kawaguchi (2011) in Japan indicate that males who entered primary school later due to school entry laws accumulated 13.20 years of education while their younger counterparts only had 13.03 years of education. Moreover, Valdés and Requena (2024) find that in Spain men who started school later gained 0.2 more years of schooling, on average, than their younger peers.

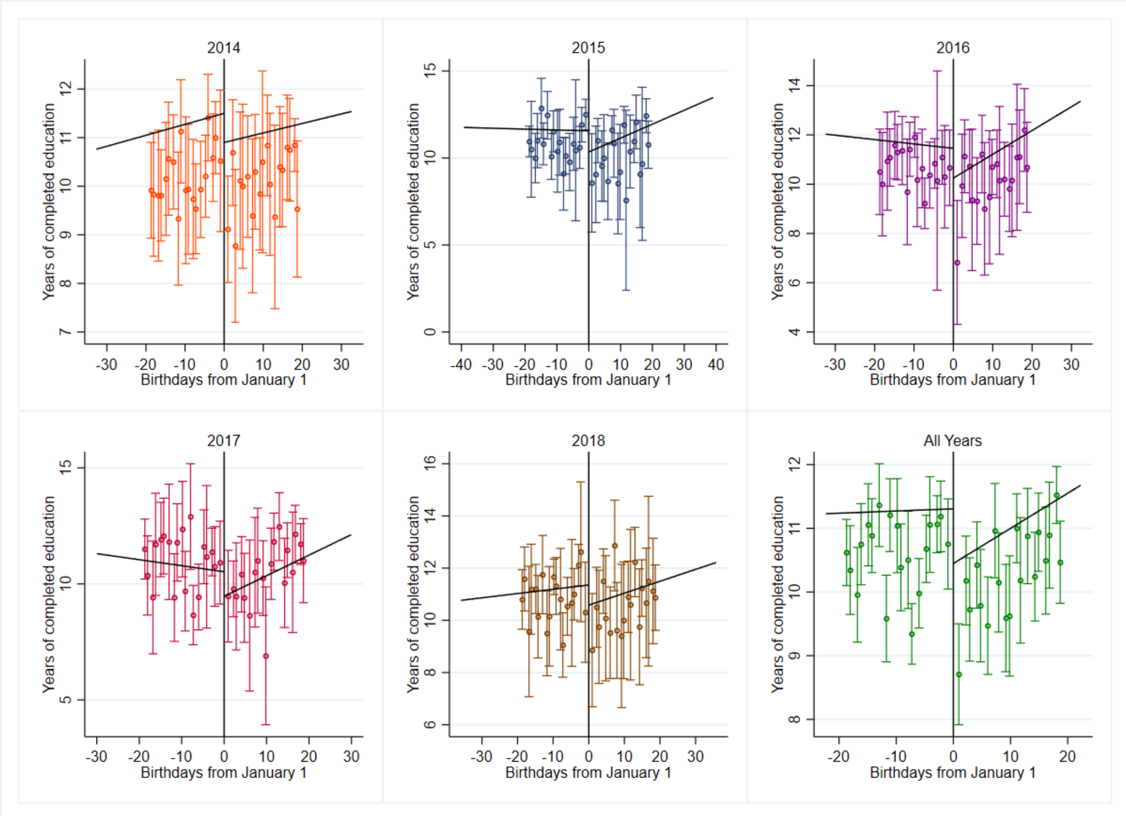
The reason for this advantage among older school entrants is that they benefit from an extra year of maturation and skill development (Valdés and Requena, 2024). Although younger cohorts start school earlier the higher repetition rates among them results in reduced years of schooling (Matta et al., 2016). In South Africa, van der Berg et al. (2025) find that children who start school at a younger age within the typical entry range tend to experience higher repetition rates. For example, in the Eastern Cape, the repetition rate among boys who begin school at a younger age is as high as 31%, compared to 19% for girls in the same age group. These high repetition rates are often associated with poor academic performance, which may ultimately result in school dropout or reduced educational attainment.

Females

The impact of the school entry law on women's years of education is depicted in Figure 4.6 and Table 4.4. As with the overall sample, the said figure shows that females who started school early experience reduced years of education throughout the study period. Table 4.4 confirms this negative impact. Panel (b), which presents estimates based on the donut RD (excluding observations whose birth date falls on 1 January) indicates that the SEA significantly reduced the years of completed education by between 9 and 15 months. The estimates are considerably smaller compared to the uncorrected ones (Panel (a)) except for 2016, again emphasizing the significance of the manipulation. These results are consistent with the literature as discussed earlier.

The school entry age strongly and significantly reduces the years of education throughout the five-year period. The estimates ranges from -0.55 in 2018 to -1.2 years in 2016. These results indicate that a one-year increase in the age at which an individual started school is associated with a reduction of between 0.55 and 1.2 years of education. However, the pooled estimate is -1.1 years, suggesting some biasses in years of education especially when compared to 2014 and 2018. These results suggest that school entry age is a strong instrument for education. Moreover they align with some of the studies, which find that starting school early reduces educational

Figure 4.4: Impact of school entry age on years of completed education for women



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

attainment among women (Kawaguchi, 2011).

Males

The effect of school entry age on men's years of education is shown in Figure 4.4 and Table 4.5. The figure shows that men who started school early have lower years of education compared to those who started late. Table 4.5 confirms the negative effect. Two sets of estimates are shown: Panel (a) shows the fuzzy RD estimates that are based on all available data, including the cut-off (0), which may be subject to manipulation as seen before. Panel (b) depicts estimates that exclude the cut-off, mitigating manipulation concerns.

Panel (b) of the table under the first stage shows a negative coefficient for school entry age, which is statistically significant throughout the survey years indicating that SEA significantly affects the probability of receiving more education. The negative sign is an indication that individuals

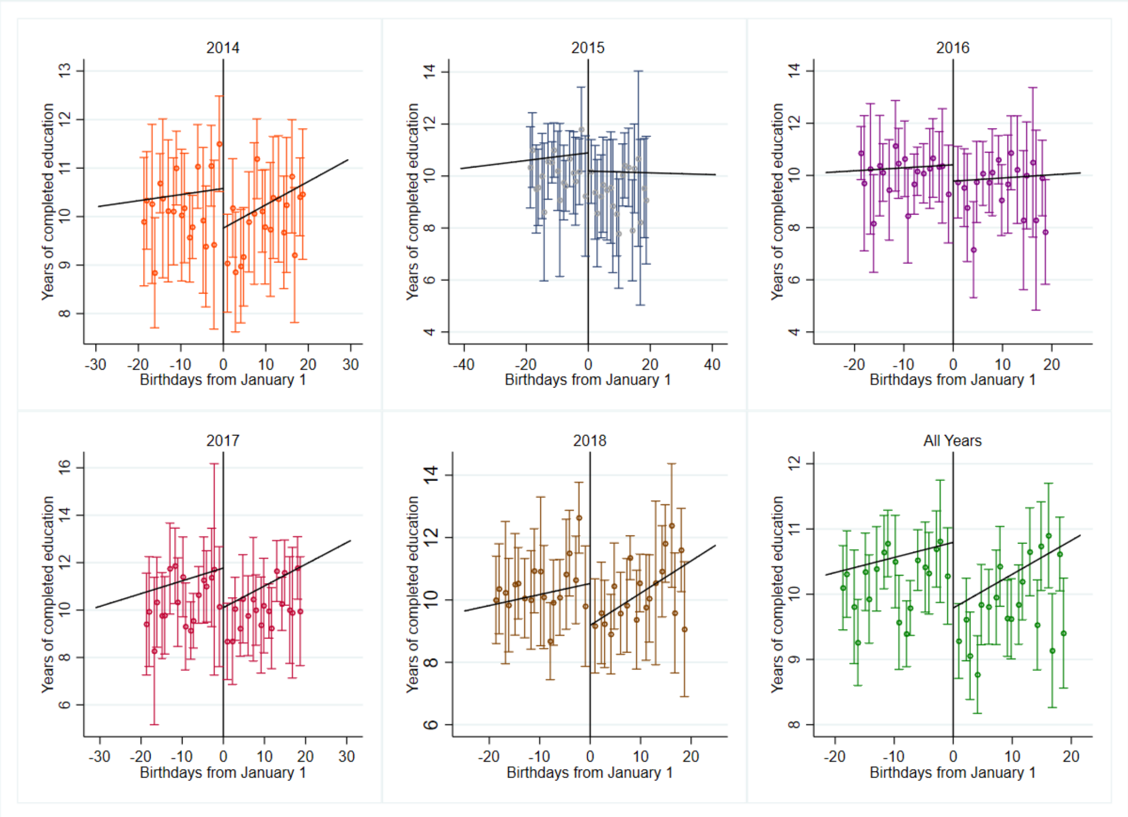
Table 4.4: Impact on years of education and the logarithm of monthly earnings for women

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
	Panel (a): Fuzzy RD with potential manipulation					
First stage						
School entry age	-0.946***	-1.233***	-1.092***	-1.317***	-1.021**	-1.486***
	(0.327)	(0.370)	(0.406)	(0.425)	(0.410)	(0.230)
Reduced Form						
School entry age	-0.192*	-0.350***	-0.222	-0.287**	-0.140	-0.337***
	(0.114)	(0.120)	(0.155)	(0.131)	(0.117)	(0.083)
Second stage						
Education	0.203*	0.284***	0.203	0.218**	0.137	0.227***
	(0.113)	(0.099)	(0.133)	(0.091)	(0.113)	(0.049)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	34.792	45.770	32.743	30.074	34.895	15.706
Bandwidth Bias (b)	83.917	105.976	77.248	73.470	80.283	52.529
	Panel (b): Fuzzy RD corrected for manipulation (donut RD)					
First stage						
School entry age	-0.772**	-0.984**	-1.230**	-1.117**	-0.475	-1.072***
	(0.393)	(0.439)	(0.487)	(0.494)	(0.466)	(0.246)
Reduced Form						
School entry age	-0.276*	-0.184	-0.273	-0.259	-0.131	0.299***
	(0.152)	(0.143)	(0.183)	(0.162)	(0.147)	(0.082)
Second stage						
Education	0.357	0.188	0.222	0.232*	0.275	0.299***
	(0.222)	(0.133)	(0.141)	(0.129)	(0.374)	(0.082)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	31.989	42.519	32.339	28.945	34.006	22.230
Bandwidth Bias (b)	74.071	105.926	84.851	72.958	81.391	62.191

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Bandwidths below 4 were not feasible due to few observations. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*). Bandwidths below 4 were not feasible due to few observations.

who started school earlier (younger cohort) have lower educational attainment compared to those who enrolled later. The corrected coefficients are smaller in magnitude compared to the uncorrected ones, which indicates the extent of the manipulation effect. However, the coefficients remain statistically significant throughout the survey years. It is evident that school entry age had the smallest impact on years of education in 2015 with a coefficient of -0.692. This estimate shows that starting school one day earlier significantly reduces education years by about 8 months. The impact is larger in the other years, with the largest effect in 2017 in which a one-day earlier entry into school is associated with a reduction of of 17 months of education. These results are consistent with the literature, which finds that individuals who start school later end up benefiting more in terms of educational attainment compared to those who starts school earlier. In Spain, for example, Valdés and Requena (2024) find that males who started school later (older

Figure 4.5: Impact of school entry age on years of completed education for men



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.4. Source: Author's calculations based on data from the GHS 2014 and 2016.

cohort) accumulate 0.2 more schooling years than males who started school earlier (younger cohort). In South Africa, as discussed before, van der Berg et al. (2025) find that children who start school at a younger age tend to experience higher repetition rates. For instance, in the Eastern Cape, the repetition rate among boys who begin school at a younger age is as high as 31%. These elevated repetition rates are often linked to poor academic performance, which may ultimately lead to school dropout or reduced educational attainment.

4.4.4 Impact on earnings

Having established that school entry age (SEA) is a strong instrument for education, I examine the effect of years of education on labour market earnings. Before I do that I discuss the reduced-form effects. As stated earlier, most of the literature assesses the impact of school entry age on earnings using reduced-form estimates. The reduced form captures how being just eligible for

Table 4.5: Impact on years of education and logarithm of monthly earnings for men

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
	Panel (a): Fuzzy RD with potential manipulation					
First stage						
School entry as	-1.379***	-1.148 ***	-1.214 ***	-2.211***	-1.737 ***	-1.538***
	(0.337)	(0.373)	(0.381)	(0.481)	(0.422)	(0.182)
Reduced form						
School entry age	-0.318**	-0.105	-0.125	-0.343**	0.060	-0.167***
	(0.145)	(0.117)	(0.123)	(0.137)	(0.123)	(0.060)
Second stage						
Education	0.231**	0.092	0.103	0.155***	-0.035	0.109***
	(0.101)	(0.092)	(0.098)	(0.057)	(0.076)	(0.037)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	29.528	33.837	27.839	26.434	26.228	26.933
Bandwidth Bias (b)	71.478	86.607	72.559	73.731	61.135	80.935
	Panel (b): Fuzzy RD corrected for manipulation (donut RD)					
First stage						
School entry age	-1.078***	-0.692*	-0.785*	-1.427***	-1.376***	-1.196***
	(0.394)	(0.407)	(0.469)	(0.529)	(0.495)	(0.238)
Reduced form						
School entry age	-0.255*	0.008	-0.082	-0.179	0.140	-0.077
	(0.154)	(0.127)	(0.166)	(0.140)	(0.141)	(0.077)
Second stage						
Education	0.238	-0.013	0.104	0.125	-0.102	0.064
	(0.148)	(0.188)	(0.207)	(0.092)	(0.124)	(0.063)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	30.774	38.473	25.270	28.770	24.783	21.542
Bandwidth Bias (b)	74.554	100.141	67.219	70.575	60.768	73.707

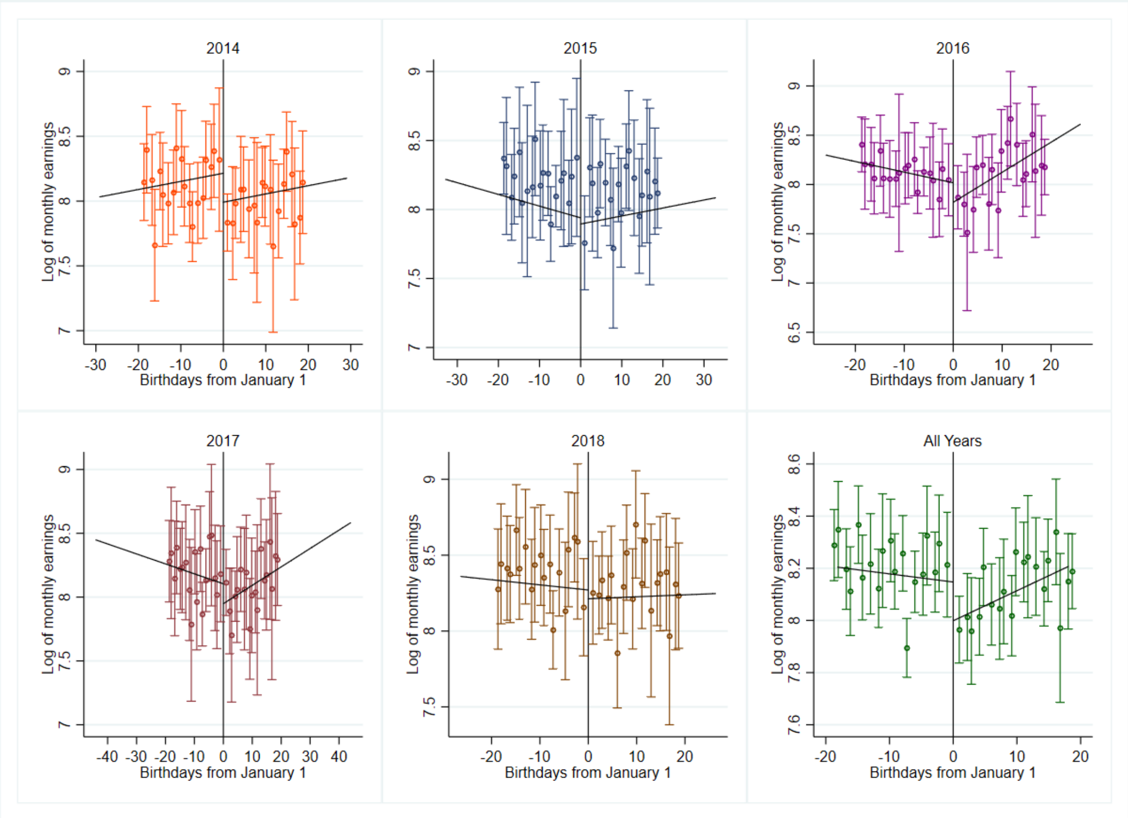
Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

early school entry affects labour market earnings regardless of whether individuals actually obtained more education. As with education, I include control variables such as marital status, age group, province and sex. These covariates show evidence of discontinuities at the cut-off (see Appendix C), and are therefore included as control variables in the analysis. As before, I discuss the overall sample estimates before I split the discussion by gender.

Overall Sample

The reduced form impact of SEA on earnings for the overall sample is presented in Figure 4.5 and Table 4.3. The vertical axis in the graph represents the logarithm of real monthly earnings, while the horizontal axis corresponds to the number of birthdays, which is the running variable centered around January 1. The vertical line at zero divides the sample into two groups: the control group (to the left of zero) and the treated group (to the right of zero, including zero). The solid curves on either side of zero show the regression functions estimated using the 'rdplot' command in Stata. Figure 4.5 shows a clear downward jump in log monthly earnings at the

Figure 4.6: Impact of school entry age on the logarithm of monthly earnings for the overall sample



Note: The graphs are based on the optimal bandwidths estimated in Table 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

cut-off in all survey years, indicating that individuals who started school earlier (younger cohort) earn significantly less as adults compared to those who started school later (older cohort). This discontinuity suggests a causal effect of age at school entry on adult earnings. The downward jump in log monthly earnings is confirmed by the negative sign of the coefficient on school entry age in the reduced-form equation in Panel (b) of Table 4.3. However, SEA is statistically significant only in 2014, 2016 and 2017 while the pooled estimate is also significant. Accordingly, early school entry is associated with 24.6, 21.7 and 15.6% reduction in monthly earnings in the respective years. The pooled estimate is 22.5%, on average.

The result aligns with the findings in the literature. Peña (2017) finds that in Mexico, the younger cohort earn 1.9% (men) and 3.8% (women) less than their older counterparts. In Brazil, the impact is even larger. Matta et al. (2016) find that the earnings of the older cohort is between 15% and 22% higher than those of the younger cohort. Pehkonen et al. (2015) and Larsen and Solli (2017) find similar results for Finland and Norway, respectively. Kawaguchi (2011) finds that older children take advantage of their maturation at school entry to obtain more education years than their younger counterparts, which eventually leads to higher earnings.

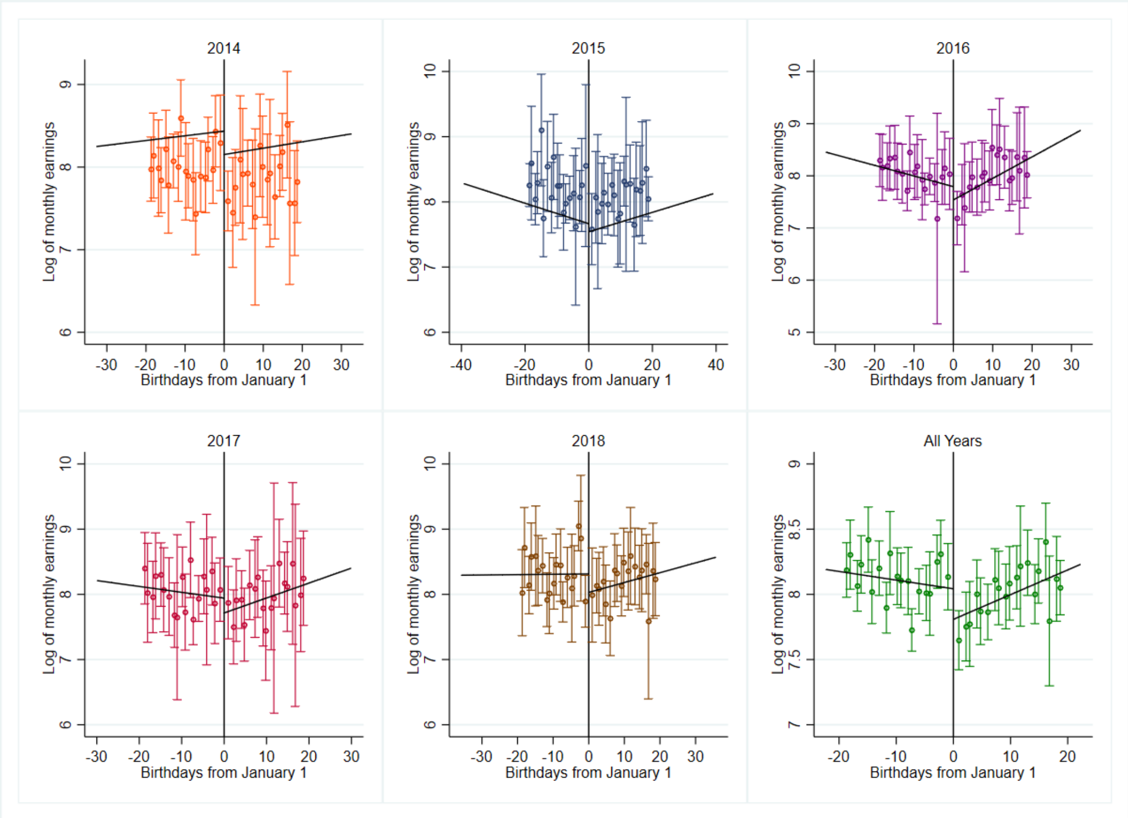
However, the reduced form effects are limited to the effect of the SEA on earnings, they reflect the intention to treat regardless of whether actual treatment was received or not. Hence, they do not tell us whether those who eventually started school early benefited in terms of earnings. To investigate the effect on education on earnings, I turn to the second-stage estimates and assess whether the earnings of those who started school early increased or not. We have seen already that SEA is a strong instrument for education. Panel (b) of Table 4.3 shows the estimates for the corrected version of the RD model. The coefficient on education is positive and statistically significant in 2014, 2016 and 2017, but larger in magnitude compared to the uncorrected estimates (see Panel (a)) of the table. The pooled estimate is also positive, statistically significant and larger.

The effect of education on earnings although positive and statistically significant is heterogeneous across the years, with 2014 experiencing higher returns of 30.2% compared to 20.9 and 20.3% in 2016 and 2017, respectively. Therefore, an additional year of education attained among early school entrants increased their monthly earnings by between 20.3% and 30.2%, on average. Next, I examine the impact on earnings by gender.

Females

Two effects are examined here. The effect of the instrument on earnings (reduced-form estimates) and the impact of education on earnings (second stage). Figure 4.7 indicates that the effect of school entry age on years of education is negative throughout the years. The size of the effect is heterogeneous across the years with 2015 having the smallest effect and 2014 the largest. Table 4.4 provides the estimates. Panel (a) shows the estimates with potential manipulation of birth dates while Panel (b) depicts results corrected for manipulation. As seen before, the difference between the two versions of estimates is evident. The RD version without correction indicates that all coefficients are negative, with statistical significance in 2014, 2015, 2017, and the pooled sample. While the estimates corrected for manipulation are also negative they are statistically

Figure 4.7: Impact of school entry age on the logarithm of monthly earnings for women



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014-2018.

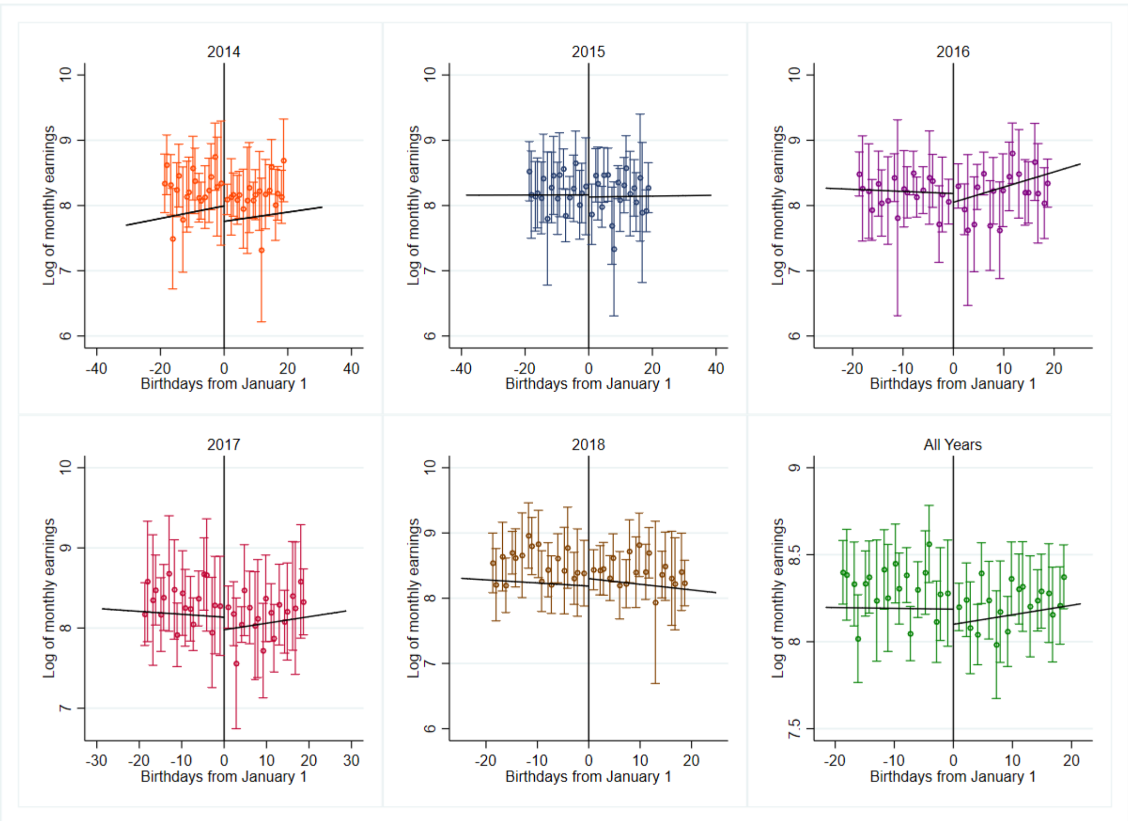
significant only in 2014 and for the pooled sample, which underscores the importance of the correction. The coefficient of -0.276 for 2014 suggests that early school entry reduces earnings by about 28% among women, on average. This finding is in line with several studies. For example, Liu and Marois (2023) find that starting school later significantly increases annual earnings by 59% for non-agricultural jobs in China. In Mexico, however, the size of the effect is relatively smaller with Peña (2017) finding that the earnings of women who started school early are 3.8% lower than for late school entrants.

Turning to the impact of education on earnings, it is evident that the corrected estimates differ from the uncorrected ones. Focusing on the corrected version, one sees that the effect of education on earnings is positive and heterogeneous across the years. The coefficient on education is statistically significant in 2017 with the treatment effect of 0.232, respectively. Therefore, an additional year of education increased the monthly earnings of early school entrants by 23.2% in 2017.

Males

As with the overall and female samples, I analyse the effect on earnings in two ways. The first is to look at the reduced-form estimates, which represents the intention to treat effect. This is the most common way the literature examines the effect of school entry age on labour market earnings. However, I go beyond that and estimate the causal effect of education on earnings using school entry age as an instrument for education. Thus, the second way I examine the impact on earnings is to estimate the returns to education among individuals who actually started school early due to the school entry rule. In the first analysis, I rely on both figure 4.8 and Table 4.5.

Figure 4.8: Impact of school entry age on the logarithm of monthly earnings for men



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.4. Source: Author's calculations based on data from the GHS 2014-2018.

The figure appears to indicate that school entry age had a negative effect on years of education in 2014, 2016 and 2017. However, on a closer examination, Table 4.5 shows that the effect of SEA is statistically significant only in 2014. The uncorrected version shows statistical significance in 2017 as well. Thus, the coefficient of -0.255 in 2014 suggest that starting school earlier was associated with a reduction of 25.5% in monthly earnings in 2014 compared to starting school later. As noted before, this result is consistent with findings in other countries. For example, Peña (2017) find than younger cohorts of men earn 1.9% lower than the older cohort in Mexico.

Moving on to the second analysis, the causal effect of education on earnings is positive in 2014, 2016 and 2017 and negative in the other years, but statistically insignificant in all the years including the pooled sample. Studies such as Dobkin and Ferreira (2010) and Pehkonen et al. (2015) find similar results even though their estimates are based on reduced-form effects.

4.4.5 Robustness checks

I conduct a range of robustness checks to ensure that the findings are reliable and not the result of spurious relationships in the data. The two robustness checks that I consider are (1) estimating the treatment effect at placebo cut-off points of the running variable, and (2) checking for balance on pre-determined covariates.

Placebo cut-off points

The analysis of placebo cut-off points is a robustness check used to validate the credibility of the RD design. Specifically, it involves testing for discontinuities at arbitrary points along the running variable, points at which no treatment should occur. The absence of discontinuities at these placebo thresholds provides support for the claim that the observed discontinuity at the true cut-off is attributable to the treatment. Conversely, if significant discontinuities are detected at these artificial thresholds, it raises concerns about the internal validity of the RD design, suggesting that the observed effects may be driven by other unobserved factors correlated with the running variable. In this analysis, I conduct placebo tests by estimating the RD at eight alternative cut-off points, four to the left and four to the right of the actual threshold. The results are presented in Table 4.6.

Table 4.6: Impact of education on the logarithm of monthly earnings at different placebo cut-offs for the overall sample

	(1) c=-40	(2) c=-30	(3) c=-20	(4) c=-10	(5) c=10	(6) c=20	(7) c=30	(8) c=40
				Panel (a): Overall Sample				
2014	0.448	-1.079	-0.644	0.435	0.111	-1.410	0.152	-0.101
	(0.343)	(3.024)	(14.290)	(0.877)	(0.337)	(5.273)	(0.159)	(0.391)
2015	0.108	0.157	0.310	0.860	0.210	-0.442	0.042	-0.113
	(0.193)	(0.115)	(0.485)	(1.020)	(0.218)	(2.468)	(0.243)	(0.553)
2016	1.099	0.002	-0.349	0.282	1.764	-0.279	0.091	0.804
	(1.189)	(0.147)	(2.004)	(0.321)	(1.961)	(0.320)	(0.137)	(1.135)
2017	-0.306	0.318	-0.201	0.042	0.138	-0.060	6.158	0.521
	(0.969)	(0.707)	(0.778)	(0.254)	(0.155)	(0.350)	(53.454)	(0.407)
2018	0.379**	0.249	-1.118	3.646	0.224	-0.100	0.351*	-1.640
	(0.190)	(3.792)	(7.105)	(31.107)	(0.173)	(0.285)	(0.201)	(26.104)
All Years	0.612	0.101	0.473	0.365	0.479*	0.077	0.251**	-0.021
	(0.636)	(0.191)	(2.783)	(0.322)	(0.269)	(0.293)	(0.109)	(0.302)
				Panel (b): Female Sample				
2014	0.396*	-0.169	0.560	-1.452	0.072	-1.311	0.143	0.187
	(0.224)	(0.743)	(2.714)	(5.788)	(0.561)	(100.961)	(0.136)	(0.365)
2015	4.636	0.242*	0.247*	0.406*	0.127	0.131	0.388	0.078
	(28.839)	(0.131)	(0.133)	(0.240)	(0.455)	(0.110)	(0.733)	(0.229)
2016	0.750	0.209	0.083	0.657	0.656**	0.481	0.367	0.433*
	(0.561)	(0.162)	(0.443)	(0.535)	(0.319)	(0.506)	(0.274)	(0.229)
2017	-2.536	0.546	0.203	-0.233	0.170	-1.817	0.499	-0.629
	(9.843)	(0.549)	(0.337)	(0.730)	(0.227)	(20.595)	(0.655)	(2.883)
2018	0.274	0.340	-0.168	-0.185	0.784	0.280	-0.015	-0.505
	(0.208)	(0.878)	(1.089)	(1.083)	(1.496)	(0.584)	(0.797)	(1.302)
All Years	3.475	0.111	0.250	0.777	0.404**	0.174	0.183	0.141
	(41.812)	(0.184)	(0.378)	(0.574)	(0.169)	(0.214)	(0.128)	(0.488)
				Panel (c): Male Sample				
2014	0.414	-1.449	-0.076	-0.215	0.162	-0.723	0.230	-0.394
	(0.833)	(4.838)	(1.372)	(0.465)	(0.453)	(1.647)	(0.242)	(0.935)
2015	0.038	0.070	0.455	-0.593	0.206	0.091	0.057	0.072
	(0.152)	(0.256)	(0.974)	(88.953)	(0.378)	(0.131)	(0.146)	(0.243)
2016	1.122	-0.124	0.550	-0.577	-1.634	-0.151	0.039	0.196
	(2.062)	(0.228)	(0.529)	(1.852)	(338.023)	(0.156)	(0.250)	(0.201)
2017	0.315	0.723	-0.270	-0.118	0.189	0.023	2.979	1.021
	(0.651)	(15.312)	(0.752)	(0.598)	(0.202)	(0.167)	(12.919)	(1.813)
2018	1.106	-0.081	-0.032	1.169	0.024	-0.356	0.302**	-0.286
	(5.801)	(0.506)	(0.874)	(3.412)	(0.161)	(0.523)	(0.149)	(1.047)
All Years	-0.073	-0.021	0.298	-0.043	0.330	0.352	0.238*	0.819
	(0.934)	(0.324)	(0.438)	(0.390)	(0.238)	(0.437)	(0.122)	(2.030)

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

The premise of the placebo test is that if treatment is truly assigned based on the January 1 cut-off, then no significant discontinuities should be observed at any other point along the running variable. The results confirm this expectation for most cases. For the overall sample, no statistically significant discontinuities are found in 2014 through 2017. However, two significant jumps are observed in 2018, one on either side of the true cut-off. In the pooled sample, two discontinuities appear on the right side of the threshold.

When disaggregated by gender, the female sample shows some evidence of discontinuity: one in 2014, three in 2015, two in 2016, and one in the pooled sample. By contrast, the male sample exhibits no statistically significant discontinuities at any placebo cut-off point. Overall, while a few significant discontinuities emerge, primarily within the female subsample, most placebo estimates are statistically insignificant. This pattern provides reassurance that the RD estimates are not driven by arbitrary breaks in the running variable, thereby supporting the internal validity of the identification strategy.

Covariate balance

Balance on covariates involves examining whether the observable characteristics of units just below and just above the threshold are similar. If these covariates are balanced, this supports the idea that units near the threshold are comparable, reinforcing the validity of the RD design. Tables 4.7, 4.8 and 4.9 show the impact of school entry age on the pre-determined characteristics of units. Statistically significant differences in units around the threshold would call the RD design into question.

Table 4.7 depicts the impact of school entry age on language of interview for the overall and female samples. The respondents were interviewed in eleven languages as shown in the table. If the RD design is valid, school entry age should not systematically influence the language of interview, i.e., the coefficients should be small and statistically insignificant. Most of the coefficients are small in magnitude and statistically insignificant, indicating no systematic relationship between school entry age and language of interview. This supports the validity of the RD design.

Table 4.8 presents the results of covariate balance tests evaluating whether the school entry age cut-off affects the language of interview for the male sample across the survey years 2014 to 2018, as well as for the pooled sample. For the majority of years and language groups, the estimated coefficients are small in magnitude and statistically insignificant, suggesting that the running variable is generally balanced with respect to this covariate, a key requirement for the validity of the RD design.

Nonetheless, a few notable exceptions are observed. In 2014, Afrikaans shows a modest but statistically significant negative coefficient. In 2015, Sesotho and Siswati exhibit significant discontinuities. For 2016, Setswana and Zulu show positive and statistically significant differences.

Table 4.7: Covariate balance: assessing impact of school entry age on language of interview for the overall and female samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample						
Afrikaans	-0.024**	-0.007	0.008	0.005	0.004	-0.000
	(0.012)	(0.008)	(0.010)	(0.011)	(0.006)	(0.004)
English	0.010	-0.004	-0.004	0.016	-0.009	0.002
	(0.012)	(0.022)	(0.022)	(0.024)	(0.021)	(0.008)
Ndebele	-0.007	0.014	0.007	0.026***	0.006	0.007
	(0.007)	(0.009)	(0.009)	(0.010)	(0.008)	(0.004)
Sepedi	0.009	-0.010	0.002	0.063***	0.004	0.014
	(0.017)	(0.023)	(0.020)	(0.024)	(0.019)	(0.010)
Sesotho	-0.047**	-0.026	-0.021	-0.032	0.017	-0.022**
	(0.020)	(0.020)	(0.019)	(0.020)	(0.019)	(0.009)
Siswati	0.020*	-0.004	-0.008	-0.028	-0.017	-0.004
	(0.012)	(0.011)	(0.011)	(0.019)	(0.016)	(0.007)
Tsonga	-0.007	0.019	0.007	0.014	0.014	0.007
	(0.009)	(0.018)	(0.016)	(0.012)	(0.012)	(0.006)
Setswana	0.014	0.045**	-0.026	-0.069***	-0.014	-0.005
	(0.015)	(0.020)	(0.021)	(0.022)	(0.021)	(0.009)
Tshivenda	0.000	-0.016	-0.005	0.007	-0.001	-0.005
	(0.006)	(0.012)	(0.010)	(0.016)	(0.013)	(0.005)
Xhosa	0.019*	0.003	-0.000	-0.013	-0.006	0.004
	(0.011)	(0.017)	(0.014)	(0.017)	(0.012)	(0.007)
Zulu	-0.007	0.004	0.048**	-0.011	-0.015	0.005
	(0.018)	(0.030)	(0.022)	(0.028)	(0.029)	(0.009)
Panel (b): Female Sample						
Afrikaans	-0.011	-0.027*	0.009	0.016	0.009	0.001
	(0.014)	(0.015)	(0.016)	(0.011)	(0.008)	(0.006)
English	0.024	-0.075**	0.001	-0.022	-0.056**	-0.016
	(0.020)	(0.030)	(0.029)	(0.030)	(0.028)	(0.012)
Ndebele	-0.014	0.030*	0.010	0.020	0.005	0.009
	(0.009)	(0.016)	(0.013)	(0.017)	(0.013)	(0.007)
Sepedi	0.024	-0.006	0.018	0.025	-0.005	0.007
	(0.023)	(0.037)	(0.033)	(0.031)	(0.027)	(0.012)
Sesotho	-0.037	0.008	-0.028	0.002	0.053*	-0.005
	(0.025)	(0.029)	(0.031)	(0.027)	(0.028)	(0.012)
Siswati	0.031**	-0.049**	-0.004	-0.049*	-0.053***	-0.016
	(0.015)	(0.022)	(0.022)	(0.028)	(0.020)	(0.010)
Tsonga	-0.003	0.001	0.019	0.010	0.051**	0.010
	(0.014)	(0.017)	(0.020)	(0.019)	(0.022)	(0.008)
Setswana	0.013	0.107***	0.005	-0.028	-0.018	0.013
	(0.021)	(0.037)	(0.027)	(0.027)	(0.026)	(0.012)
Tshivenda	0.003	-0.021	-0.017	-0.015	-0.004	-0.010
	(0.010)	(0.016)	(0.014)	(0.023)	(0.018)	(0.007)
Xhosa	0.002	0.007	-0.019	-0.023	0.003	-0.004
	(0.015)	(0.021)	(0.022)	(0.026)	(0.017)	(0.009)
Zulu	-0.034	0.029	0.014	0.070	0.008	0.008
	(0.026)	(0.037)	(0.041)	(0.042)	(0.035)	(0.014)

Note: The dependent variable is the the respective language while the independent variable is the running variable (birthdays). Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant languages were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

In 2017, significant imbalances are detected for Ndebele, Sesotho, and Setswana, while Sepedi shows a significant increase. Finally, in the pooled sample, only Sesotho is associated with a statistically significant discontinuity.

Overall, although a handful of statistically significant differences are detected in specific years and languages, the general pattern supports covariate balance on the language of interview for the male sample, thereby reinforcing the validity of the RD design.

Table 4.9 tests whether school entry age assignment is balanced across different household

Table 4.8: Covariate balance: assessing impact of school entry age on language of interview for the male sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Afrikaans	-0.027*	0.010	0.008	-0.010	0.012	-0.008
	(0.016)	(0.010)	(0.012)	(0.015)	(0.011)	(0.007)
English	0.002	0.038	-0.002	0.034	0.031	0.017
	(0.019)	(0.034)	(0.031)	(0.032)	(0.027)	(0.011)
Ndebele	-0.001	0.002	-0.003	0.027***	0.002	0.005
	(0.009)	(0.010)	(0.010)	(0.010)	(0.008)	(0.005)
Sepedi	-0.004	-0.019	-0.016	0.103***	0.010	0.013
	(0.022)	(0.036)	(0.030)	(0.036)	(0.025)	(0.015)
Sesotho	-0.038	-0.069**	-0.021	-0.061**	-0.013	-0.044***
	(0.024)	(0.029)	(0.026)	(0.028)	(0.024)	(0.014)
Siswati	0.007	0.026*	-0.017	-0.011	0.013	0.005
	(0.018)	(0.015)	(0.013)	(0.023)	(0.022)	(0.009)
Tsonga	-0.014	0.023	-0.005	0.023	0.003	0.003
	(0.014)	(0.027)	(0.024)	(0.017)	(0.019)	(0.009)
Setswana	0.033	0.006	-0.052*	-0.095***	-0.005	-0.017
	(0.022)	(0.033)	(0.030)	(0.032)	(0.029)	(0.013)
Tshivenda	-0.000	-0.013	0.007	0.024	-0.002	0.003
	(0.007)	(0.016)	(0.014)	(0.021)	(0.017)	(0.007)
Xhosa	0.015	-0.005	0.012	-0.004	-0.011	0.011
	(0.014)	(0.024)	(0.022)	(0.020)	(0.021)	(0.010)
Zulu	0.021	0.013	0.081**	-0.067*	-0.031	0.003
	(0.024)	(0.044)	(0.033)	(0.040)	(0.039)	(0.014)

Note: The dependent variable is the the respective language while the independent variable is the running variable (birthdays). Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant languages were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

relationships across survey years. The goal is to ensure there are no systematic differences in household composition that could bias the RD design. Most coefficients are small and statistically insignificant, suggesting no strong relationship between school entry age and relationship to the head of household. This supports the validity of the RD design.

In general, I find that the pre-determined covariates are balanced. Where I find imbalance, I incorporate the the concerned coviates in the estimation of relevant RD models.

4.4.6 Sensitivity analysis

I conduct three sensitivity tests. These include varying the bandwidth size, using donut RD, and varying the polynomial order.

Bandwidth size

Selecting an optimal bandwidth involves a variance-bias trade-off. A narrower bandwidth yields a more precise estimate around the cut-off point, yet it risks amplifying noise. Conversely, a broader bandwidth diminishes noise but may obscure the treatment effect. Hence, careful consideration of data characteristics and the research question is vital in determining the optimal bandwidth

Table 4.9: Covariate balance: assessing impact of school entry age on relation to head of household for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
			Panel (a): Overall Sample			
Head of Household (HHD)	-0.060*	0.013	-0.017	0.032	0.042	-0.010
	(0.032)	(0.039)	(0.035)	(0.035)	(0.045)	(0.016)
Spouse	0.033	-0.026	0.011	-0.010	-0.004	-0.002
	(0.021)	(0.025)	(0.030)	(0.029)	(0.026)	(0.011)
Child	-0.026	0.017	-0.038	-0.038	-0.041	-0.024*
	(0.022)	(0.023)	(0.031)	(0.029)	(0.032)	(0.014)
Sibling	-0.008	0.004	-0.033*	-0.005	0.003	-0.010
	(0.013)	(0.016)	(0.018)	(0.013)	(0.018)	(0.006)
Parent	0.001	0.003	0.001	0.002	0.000	0.001
	(0.002)	(0.003)	(0.001)	(0.002)	(0.000)	(0.001)
Grandchild	0.015	-0.007	0.024**	0.009	0.010	0.010*
	(0.011)	(0.011)	(0.011)	(0.008)	(0.009)	(0.005)
Other relative	0.039**	0.012	0.020	0.005	0.021	0.021***
	(0.017)	(0.016)	(0.014)	(0.012)	(0.013)	(0.007)
Not related	0.002	-0.005	-0.004	-0.006	-0.003	-0.000
	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	(0.003)
			Panel (b): Female Sample			
Head of Household	-0.042	0.017	-0.001	0.029	0.009	-0.016
	(0.035)	(0.061)	(0.054)	(0.057)	(0.060)	(0.026)
Spouse	0.065*	-0.061	-0.035	-0.054	-0.054	-0.031*
	(0.034)	(0.040)	(0.053)	(0.051)	(0.044)	(0.016)
Child	-0.028	0.005	0.015	0.020	0.026	-0.000
	(0.034)	(0.042)	(0.043)	(0.044)	(0.042)	(0.020)
Sibling	-0.000	-0.011	-0.036	-0.006	-0.011	-0.016
	(0.014)	(0.021)	(0.025)	(0.017)	(0.019)	(0.010)
Parent	0.003	0.007	0.002	0.003	0.000	0.002
	(0.003)	(0.008)	(0.003)	(0.004)	(0.002)	(0.002)
Grandchild	0.029*	-0.013	0.011	0.018	0.016	0.016**
	(0.015)	(0.017)	(0.015)	(0.015)	(0.015)	(0.007)
Other relative	0.052**	0.026	0.016	0.012	0.019	0.026**
	(0.026)	(0.023)	(0.024)	(0.020)	(0.023)	(0.012)
Not related	-0.005	-0.007	-0.003	-0.008	0.002	-0.003
	(0.007)	(0.005)	(0.011)	(0.011)	(0.007)	(0.005)
			Panel (b): Male Sample			
Head of Household	-0.002	-0.007	0.021	0.075*	0.089*	0.016
	(0.040)	(0.045)	(0.045)	(0.043)	(0.052)	(0.019)
Spouse	-0.019	-0.007	-0.008	-0.011	-0.008	-0.008
	(0.016)	(0.017)	(0.015)	(0.016)	(0.016)	(0.006)
Child	-0.000	0.017	-0.034	-0.082**	-0.109**	-0.025
	(0.032)	(0.031)	(0.034)	(0.040)	(0.044)	(0.017)
Sibling	-0.018	0.010	-0.030	0.008	0.014	-0.005
	(0.021)	(0.021)	(0.021)	(0.019)	(0.027)	(0.010)
Parent	N/A	N/A	N/A	N/A	N/A	N/A
Grandchild	0.000	-0.002	0.037***	0.004	0.006	0.007
	(0.015)	(0.013)	(0.012)	(0.009)	(0.011)	(0.007)
Other relative	0.023	-0.007	0.025	0.007	0.019	0.015*
	(0.018)	(0.018)	(0.020)	(0.014)	(0.017)	(0.008)
Not related	0.023	-0.007	0.025	0.007	0.019	0.015*
	(0.018)	(0.018)	(0.020)	(0.014)	(0.017)	(0.008)

Note: The dependent variable is the the respective relationship to the head of the household while the independent variable is the running variable (birthdays). Robust standard errors are shown in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant variables were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

method (Cattaneo, Idrobo and Titiunik, 2020). Therefore, the first sensitivity test that I applied involves examining how changes in bandwidth size affect the main results estimated in Tables 4.3, 4.4 and 4.5. I re-estimate the equations for all individuals, women and men for the period 2014-2018 (including the pooled sample) using bandwidth sizes of 10 - 90 birthdays. I use a data-driven bandwidth selection method as proposed by Calonico et al. (2017). The results are presented in Table 4.10. The optimal bandwidth for 2014 is approximately 34 days, yielding a

Table 4.10: Impact of years of education on the logarithm of monthly earnings at different bandwidths

	(1) h=10	(2) h=20	(3) h=30	(4) h=40	(5) h=50	(6) h=60	(7) h=70	(8) h=80	(9) h=90
					Panel (b): Overall Sample				
2014	0.319***	0.275**	0.287**	0.354**	0.390*	0.389	0.382	0.394	0.413
	(0.112)	(0.114)	(0.126)	(0.166)	(0.225)	(0.253)	(0.281)	(0.306)	(0.316)
2015	0.398	0.121	0.102	0.100	0.125	0.127	0.116	0.115	0.115
	(0.279)	(0.145)	(0.105)	(0.108)	(0.111)	(0.104)	(0.112)	(0.124)	(0.124)
2016	0.225**	0.288**	0.189*	0.138	0.127	0.139	0.143	0.154	0.158
	(0.106)	(0.123)	(0.112)	(0.141)	(0.169)	(0.168)	(0.205)	(0.248)	(0.269)
2017	0.100	0.121*	0.157**	0.184*	0.213*	0.231*	0.244	0.235	0.223
	(0.069)	(0.072)	(0.073)	(0.095)	(0.115)	(0.127)	(0.154)	(0.160)	(0.151)
2018	0.072	0.068	0.074	0.071	0.081	0.061	-0.029	-0.192	-0.377
	(0.086)	(0.083)	(0.106)	(0.174)	(0.268)	(0.260)	(0.317)	(0.448)	(0.559)
All Years	0.204***	0.178***	0.173***	0.185***	0.197**	0.191**	0.177*	0.164	0.148
	(0.048)	(0.049)	(0.051)	(0.066)	(0.082)	(0.085)	(0.103)	(0.127)	(0.146)
					Panel (b): Female Sample				
2014	0.533***	0.333**	0.351	0.493	0.551	0.530	0.427	0.276	0.135
	(0.195)	(0.152)	(0.215)	(0.404)	(0.635)	(0.783)	(1.040)	(1.186)	(1.300)
2015	0.293*	0.109	0.156	0.175	0.214	0.219*	0.209	0.212	0.205
	(0.158)	(0.148)	(0.120)	(0.130)	(0.136)	(0.132)	(0.150)	(0.172)	(0.168)
2016	0.334***	0.352**	0.234*	0.178	0.183	0.201	0.214	0.234	0.262
	(0.109)	(0.157)	(0.138)	(0.176)	(0.202)	(0.187)	(0.220)	(0.273)	(0.321)
2017	0.289**	0.295**	0.230*	0.232	0.301	0.314	0.309	0.332	0.414
	(0.138)	(0.144)	(0.132)	(0.163)	(0.218)	(0.241)	(0.303)	(0.343)	(0.359)
2018	0.138	0.248	0.270	0.290	0.329	0.302	0.268	0.113	-0.175
	(0.190)	(0.241)	(0.320)	(0.475)	(0.658)	(0.495)	(0.632)	(1.091)	(1.501)
All Years	0.365***	0.309***	0.270***	0.280**	0.300**	0.286*	0.257	0.208	0.161
	(0.076)	(0.082)	(0.087)	(0.115)	(0.147)	(0.148)	(0.191)	(0.256)	(0.297)
					Panel (c): Male Sample				
2014	0.155	0.212	0.238	0.275*	0.308	0.333	0.400	0.528*	0.671**
	(0.134)	(0.156)	(0.149)	(0.166)	(0.209)	(0.231)	(0.265)	(0.307)	(0.339)
2015	-0.566	0.162	0.022	-0.016	-0.006	0.003	-0.009	-0.022	-0.027
	(2.629)	(0.389)	(0.188)	(0.190)	(0.184)	(0.150)	(0.147)	(0.152)	(0.152)
2016	-0.243	0.148	0.074	0.042	0.029	0.047	0.100	0.182	0.183
	(0.679)	(0.217)	(0.200)	(0.248)	(0.315)	(0.334)	(0.439)	(0.527)	(0.521)
2017	0.005	0.062	0.136	0.148	0.096	0.075	0.047	0.008	-0.038
	(0.098)	(0.090)	(0.095)	(0.141)	(0.157)	(0.179)	(0.219)	(0.239)	(0.258)
2018	-0.102	-0.091	-0.090	-0.175	-0.211	-0.264	-0.419	-0.793	-1.199*
	(0.133)	(0.108)	(0.141)	(0.318)	(0.635)	(0.534)	(0.552)	(0.597)	(0.690)
All Years	0.028	0.058	0.078	0.083	0.082	0.084	0.088	0.102	0.108
	(0.076)	(0.065)	(0.063)	(0.082)	(0.101)	(0.105)	(0.120)	(0.132)	(0.139)

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014*a*).

statistically significant treatment effect of 30.2%. At a 30-day bandwidth, the estimate is 28.7%, and it rises to 35.4% at 40 days. Although the estimates are similar across bandwidths, statistical significance weakens as the bandwidth increases, suggesting the true effect is likely captured near the optimal range.

For 2015, the optimal bandwidth is around 33 days, producing a treatment effect of 10.6%, which is statistically insignificant. Comparable results, both in magnitude and lack of significance, are observed across other bandwidth sizes. 2016 shows an optimal bandwidth of 28 days with a statistically significant effect of 20.9%. At 30 days, the estimate is 18.9%, and it increases to

28.8% at 20 days. While the estimates remain fairly stable, both the size and significance of the effect decline as the bandwidth widens.

In 2017, the optimal bandwidth is approximately 45 days, associated with a significant treatment effect of 20.3%. The estimates remain consistent at 40 days (18.4%) and 50 days (21.3%), though significance disappears once the bandwidth exceeds 70 days. The bandwidth for 2018 is optimally set at 26 days, producing an effect of 6.5%, which is not statistically significant. Estimates across other bandwidths remain similar in both size and significance. Lastly, the pooled sample (All Years) yields an optimal bandwidth of 19 days, with a statistically significant treatment effect of 18.4%. This estimate remains stable across a range of bandwidth choices, both in magnitude and significance.

For the female sample, in 2014, the optimal bandwidth is approximately 32 days with a treatment estimate of 35.10%. The estimate increases slightly to 35.70% at a bandwidth of 30 days but remains statistically insignificant. While the point estimates vary across bandwidths, they are statistically insignificant for most bandwidth sizes. For 2015 and 2016, the estimates remain stable and statistically insignificant across bandwidths, indicating robustness to bandwidth choice.

In contrast, the 2017 estimate is statistically significant and stable at both the 29-day optimal bandwidth (23.2%) and the 30-day alternative (23.0%). Notably, the effect also remains significant at 10-day and 20-day bandwidths, underscoring its robustness. In 2018, the estimates are consistent across bandwidths (27.5% at optimal and 27.0% at alternative), suggesting similar robustness. The pooled estimate is both stable and statistically significant across bandwidths (29.9% at 22-day optimal and 30.9% at 20-day alternative), highlighting a robust average treatment effect for the female sample.

For the male sample, the estimates at most alternative bandwidth sizes are statistically insignificant across the years, consistent with the main results reported in Table 4.5.

Donut RD

Donut RD involves re-estimating the fuzzy RD model while excluding observations that are very close to the cut-off date, effectively creating a "donut hole" around the threshold. The main objective is to test whether the estimated treatment effects are sensitive to observations near the cut-off (Barreca et al., 2011; Fan et al., 2018). In this study, I have already implemented this approach by excluding individuals whose date of birth falls on 1 January. The corresponding results are presented in Tables 4.3, 4.4, and 4.5 for the overall, female, and male samples, respectively. In this sub-section, I extend the robustness check by excluding observations born within one and two days of 1 January. The results are reported for the first- and second-stage equations in Tables 4.11 and 4.12. For brevity, I do not present reduced-form estimates. However, for convenience, I replicate the ±0 donut RD estimates.

Table 4.11: Impact of school entry age on years of education at different donut holes for overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
	Panel (a): Overall Sample					
±0-day donut	-0.817***	-0.984***	-1.039***	-0.771***	-1.066***	-1.221***
	(0.267)	(0.345)	(0.343)	(0.285)	(0.351)	(0.186)
±1-day donut	-0.357	-1.060***	-1.008**	-0.640**	-1.147***	-1.029***
	(0.252)	(0.354)	(0.394)	(0.304)	(0.382)	(0.181)
±2-day donut	-0.716**	-0.963**	-0.959**	-0.434	-1.002**	-1.0399***
	(0.343)	(0.438)	(0.392)	(0.316)	(0.425)	(0.218)
	Panel (b): Female Sample					
±0-day donut	-0.772**	-0.984**	-1.230**	-1.117**	-0.475	-1.072***
	(0.393)	(0.439)	(0.487)	(0.494)	(0.466)	(0.246)
±1-day donut	-0.642	-0.754*	-0.490	-1.068**	-0.412	-0.723***
	(0.458)	(0.452)	(0.494)	(0.532)	(0.451)	(0.255)
±2-day donut	-0.871	-0.670	-0.517	-0.945	-0.282	-0.710
	(0.565)	(0.528)	(0.536)	(0.618)	(0.478)	(0.282)
	Panel (c): Male Sample					
±0-day donut	-1.078***	-0.692*	-0.785*	-1.427***	-1.376***	-1.196***
	(0.394)	(0.407)	(0.469)	(0.529)	(0.495)	(0.238)
±1-day donut	-0.586	-1.132**	-1.353***	-1.219**	-1.503***	-1.336***
	(0.468)	(0.527)	(0.519)	(0.527)	(0.529)	(0.269)
±2-day donut	-0.741	-1.219**	-1.481**	-0.915*	-1.437**	-1.253***
	(0.536)	(0.565)	(0.580)	(0.536)	(0.589)	(0.278)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol ✓ denotes yes.

The first-stage estimates are presented in Table 4.11. As previously noted, for the overall sample, the coefficient on school entry age is negative and statistically significant across all survey years, including the pooled sample. These results are reported as ±0-day donut estimates in the table. I further compare them with the ±1-day and ±2-day donut estimates. As shown in the table, both the ±1-day and ±2-day estimates suggest a negative and statistically significant effect of school entry age on education in all survey years, except for 2014 and 2017, respectively.

For the female sample, the ±0-day donut estimates indicate that school entry age has a negative and statistically significant effect on education in all years except 2018, where the estimate is insignificant. The ±1-day and ±2-day donut estimates are negative throughout the years, including in the pooled sample, but are statistically significant only in 2015, 2017, and the pooled results. Despite these differences in statistical significance, the magnitude of the coefficients remains broadly consistent across specifications. The estimates for males closely mirror those of the overall sample.

I now turn to the second-stage estimates, presenting the effect of education on earnings using the ±0-day, ±1-day and ±2-day donut samples (see Table 4.12). As previously observed, for the overall sample, the ±0 donut RD specification shows that education has a statistically significant positive effect on earnings in 2014, 2016, and 2017, with no significant effects in 2015 and 2018. The pooled estimate is also positive and statistically significant. Increasing the "donut hole" by one day yields similar results: the effect remains statistically significant in 2016 and 2017, though the significance in 2014 disappears. The pooled estimate continues to be positive and statistically significant, and its magnitude is similar to the ±0 estimate.

The estimated effects are also comparable across most years, except in 2015, where the sign reverses and the estimate becomes notably smaller. When the "donut hole" is expanded to two days, the statistical significance disappears for all individual years, except for the pooled estimate, which remains significant but with a reduced effect size. Nonetheless, the magnitude of the estimates remains relatively stable across years, suggesting that the preferred model is robust to moderate changes in bandwidth around the cut-off.

For the female sample, the ±0 donut RD estimates indicate that education has a statistically significant effect on monthly log earnings in 2017, with the pooled estimate also being positive and statistically significant. Increasing the "donut hole" by one day yields similar results: the effect remains significant in 2017, and the pooled estimate continues to be positive and significant. However, when the "donut hole" is expanded to two days, statistical significance disappears,

Table 4.12: Impact of education on the logarithm of monthly earnings at different donut holes for overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
			Panel (a): Overall Sample			
±0-day donut	0.302**	0.106	0.209*	0.203*	0.065	0.184***
	(0.133)	(0.102)	(0.114)	(0.104)	(0.098)	(0.049)
±1-day donut	0.497	-0.033	0.277*	0.294*	0.113	0.178***
	(0.394)	(0.115)	(0.165)	(0.156)	(0.117)	(0.062)
±2-day donut	0.135	-0.041	0.144	0.390	0.063	0.115*
	(0.155)	(0.152)	(0.173)	(0.270)	(0.147)	(0.070)
			Panel (b): Female Sample			
±0-day donut	0.357	0.188	0.222	0.232*	0.275	0.299***
	(0.222)	(0.133)	(0.141)	(0.129)	(0.374)	(0.082)
±1-day donut	0.270	0.127	0.171	0.285*	0.880	0.317**
	(0.301)	(0.181)	(0.404)	(0.164)	(3.130)	(0.149)
±2-day donut	-0.133	0.109	-0.009	0.218	-0.306	0.129
	(0.288)	(0.241)	(0.495)	(0.197)	(65.824)	(0.148)
			Panel (c): Male Sample			
±0-day donut	0.238	-0.013	0.104	0.125	-0.102	0.064
	(0.148)	(0.188)	(0.207)	(0.092)	(0.124)	(0.063)
±1-day donut	0.523	-0.114	0.167	0.220	-0.093	0.077
	(0.415)	(0.165)	(0.156)	(0.137)	(0.131)	(0.067)
±2-day donut	0.469	-0.072	0.126	0.311	-0.101	0.086
	(0.334)	(0.158)	(0.171)	(0.215)	(0.143)	(0.074)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol ✓ denotes yes.

although the magnitude of the effect remains broadly similar. Overall, these results suggest that the estimated effect of education on female earnings is relatively stable across small changes in the exclusion window around the cut-off.

For the male sample, the ±0 donut RD estimates indicate no statistically significant effect of education on monthly log earnings in any year or in the pooled sample. The ±1-day and ±2-day donut RD specifications yield similar findings. These results suggest that the estimated effect of education on male earnings is consistently insignificant and relatively insensitive to changes in the number of observations excluded around the threshold.

Polynomial order

Gelman and Imbens (2019) suggests that higher-order polynomials should not be used in regression discontinuity. The main results presented in Tables 4.3 – 4.5 were estimated using a first-order polynomial. In Table 4.13, I increase the polynomial order to two and compare the results from both specifications. For convenience, the first-order polynomial estimates are replicated in Table 4.13. As previously observed, for the overall sample, the first-order polynomial results show a statistically significant positive effect of education on log monthly earnings in

2014, 2016, and 2017, with no significant effects in 2015 and 2018. The pooled estimate is also positive and statistically significant.

Table 4.13: Impact of education on logarithm of monthly earnings at different polynomial orders for the overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
	Panel (a): Overall Sample (±0 Donut RD)					
Polynomial Order 1	0.302**	0.106	0.209*	0.203*	0.065	0.184***
	(0.133)	(0.102)	(0.114)	(0.104)	(0.098)	(0.049)
Polynomial Order 2	0.272**	0.111	0.231**	0.087	0.056	0.193***
	(0.110)	(0.121)	(0.105)	(0.081)	(0.097)	(0.051)
	Panel (b): Female Sample (±0 Donut RD)					
Polynomial Order 1	0.357	0.188	0.222	0.232*	0.275	0.299***
	(0.222)	(0.133)	(0.141)	(0.129)	(0.374)	(0.082)
Polynomial Order 2	0.428*	0.131	0.230	0.303**	0.266	0.364***
	(0.244)	(0.137)	(0.151)	(0.147)	(0.303)	(0.088)
	Panel (c): Male Sample (±0 Donut RD)					
Polynomial Order 1	0.238	-0.013	0.104	0.125	-0.102	0.064
	(0.148)	(0.188)	(0.207)	(0.092)	(0.124)	(0.063)
Polynomial Order 2	0.203	0.032	0.093	0.124	-0.061	0.024
	(0.148)	(0.218)	(0.227)	(0.099)	(0.105)	(0.084)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol ✓ denotes yes.

Increasing the polynomial order to two yields broadly similar results: education has a statistically significant effect in 2014 and 2016, and the pooled estimate remains positive and significant. The magnitude of the estimated effects under both specifications is comparable across most years: 2014, 2015, 2016, 2018, and in the pooled sample. However, in 2017, the second-order polynomial estimate is smaller and statistically insignificant, contrasting with the significant first-order result. Overall, the findings suggest that the estimated treatment effect of education on earnings is largely robust to the choice of polynomial order across survey years.

The first-order polynomial results indicate that education has a statistically significant positive effect on monthly log earnings for females in 2017, as well as in the pooled sample. When the polynomial order is increased to two, the estimates remain statistically significant in 2017 and the pooled sample, and additionally become significant in 2014. The estimated effect sizes are broadly similar across the two specifications. These results suggest that the impact of education on female earnings is both positive and relatively stable across different model specifications. The first-order polynomial results for males indicate that education has no statistically significant effect on monthly log earnings in any year or in the pooled sample. The second-order polynomial yields similar findings. These results suggest that education does not have a statistically significant impact on male earnings, regardless of the polynomial specification or survey year.

4.5 Conclusion

The admission policy for first time school entry into South African schools requires that children turn seven years of age by 30 June for enrollment in Grade 1, otherwise they should start shool the following year in which they turn eight years of age. The study exploits exogenous variation in educational attainment arising from the school entry policy to address the endogeneity of education and estimate the causal returns to schooling. The analysis uses data from the General Household Survey for the years 2014–2018, which includes exact information on individuals' dates of birth.

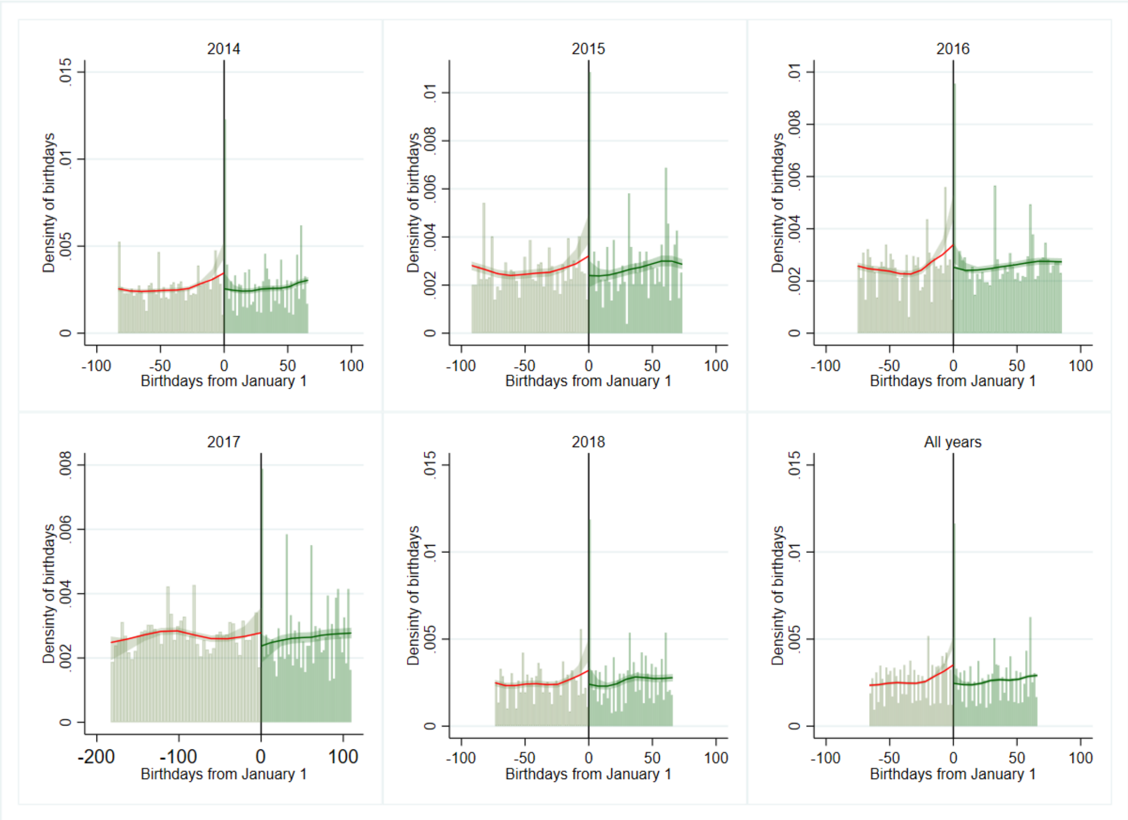
The evidence presented in this chapter indicates that there is a relatively high concentration of births around 1 January. This fact puts the RD design into question. The uneven distribution of date of birth around 1 January suggests that parents act strategically to take advantage of school entry rules. In this case parents assume that there advantages of having their children start school early. To address the problem of manipulation, I employ a donut fuzzy regression discontinuity. While numerous studies globally have investigated the impact of school starting age on labour market earnings, similar analyses remain scarce in Sub-Saharan Africa. As a result, the effects of early school entry in these contexts are largely unknown.

This study seeks to address this gap by focusing on Black South Africans, a population whose socioeconomic and institutional experiences were profoundly shaped by apartheid. The legacy of racial discrimination permeated all areas of life, including education and labour markets, relegating the Black majority to the economic margins. Although the democratic transition in 1994 abolished openly discriminatory laws, many Black South Africans continue to experience adverse socioeconomic outcomes, including poor-quality education and pervasive poverty. Most Black students still attend formerly disadvantaged schools, which often fail to deliver adequate cognitive skills. Furthermore, Black people continue to experience particularly poor labour market outcomes, such as elevated unemployment rates and disproportionately low-paying jobs among women, reflecting both historical and ongoing gender and racial inequality.

The results reveal that early school entry reduced the total number of years of education and monthly earnings, on average. However, the findings also show that among individuals who started school earlier, education significantly increased monthly earnings. These findings raise important questions about the optimal age for school entry. Early school entry could lead to higher repetition rates and poor academic performance among children who are not developmentally ready for school. Conversely, delaying school entry may improve school readiness and long-term outcomes, but it may delay entry into the labour market. Therefore, finding the right balance is critical to ensuring that children begin formal education at an age that maximizes both educational success and long-term economic outcomes.

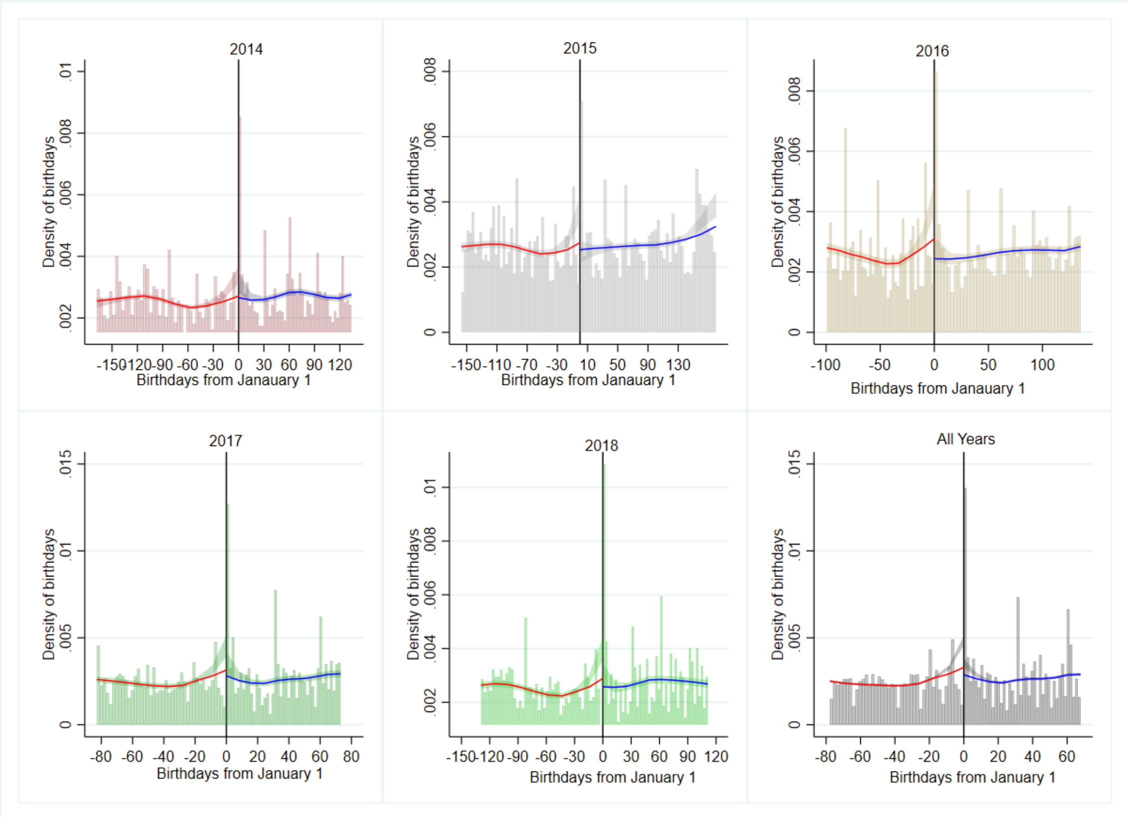
APPENDIX C

Figure C1: McCrarry manipulation test for women: 2014-2018



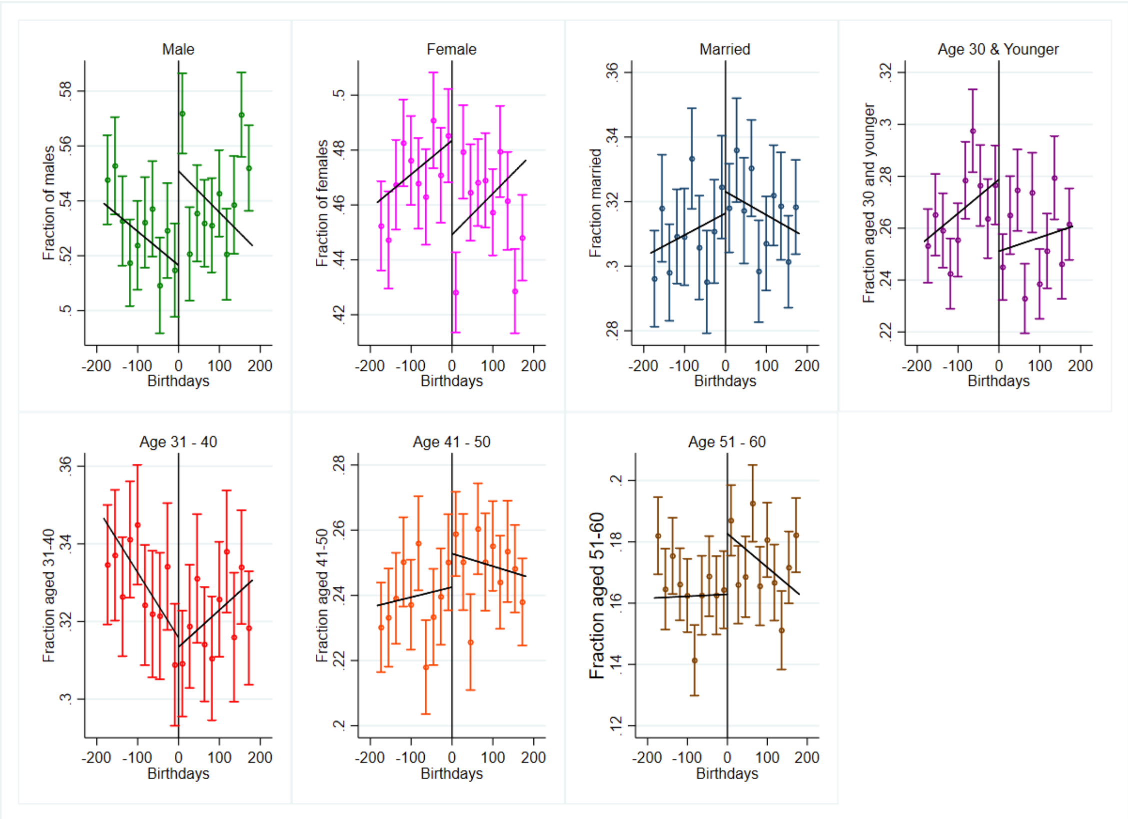
Source: Author's calculations based on data from the GHS 2014-2018.

Figure C2: McCrarry manipulation test for men: 2014-2018



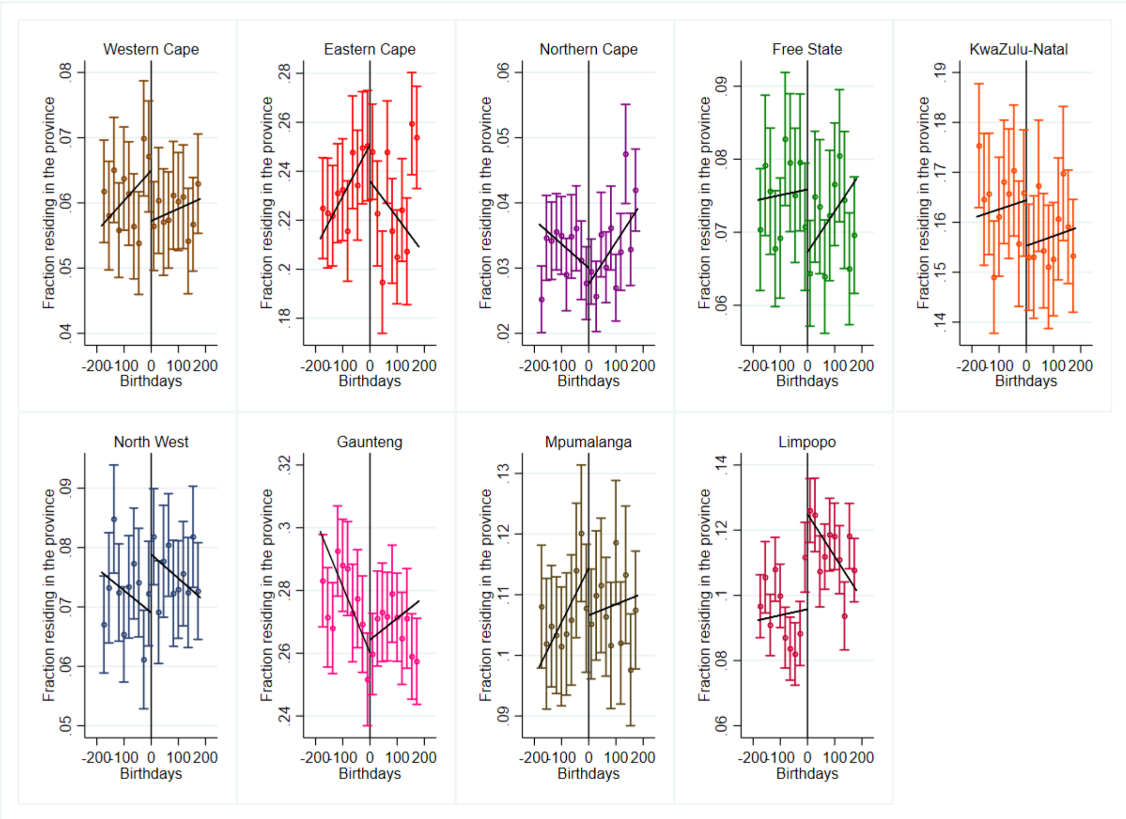
Source: Author's calculations based on data from the GHS 2014-2018.

Figure C3: Covariate balance tests: gender, marital status and age group



Source: Author's calculations based on data from the GHS 2014-2018.

Figure C4: Covariate balance tests: province of residence at the time of the survey



Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER 5

CONCLUSION

The apartheid regime in South Africa affected all spheres of life including the education system and the labour market (Mtapuri and Tinarwo, 2021). Through the racial discriminatory policies and practices of the apartheid era Black people were subjected to a separate and poor-quality education system that was meant to impart only basic skills needed for the manual labour and menial jobs that the government deemed suitable for those of their race (Thobejane, 2013; Mtapuri and Tinarwo, 2021). For the majority Black population, this meant limited career development and confinement to low-paying jobs. These policies and practices increased the income inequality between Blacks and Whites. Black women were worse off than their fellow Black men as unfavourable patriarchal beliefs and practices of the past limited their career development as they were confined to rural areas or worked in low-paying jobs such as domestic work in urban centres (Espi et al., 2019; Mosomi, 2019*b*).

The democratic transition of 1994 marked a turning point in South Africa's history, as the new government sought to dismantle apartheid-era structures and create a more inclusive society. Central to these efforts were reforms in education and the labour market. The 1996 Constitution, which enshrined equality and human rights, provided the legal foundation for these changes (Republic of South Africa, 1996*a*). One of the first major reforms was the introduction of the

Schools Act in 1996, which established a new national curriculum aimed at promoting equity in education. This act also introduced a compulsory schooling policy, ensuring that children of all races had access to basic education (Republic of South Africa, 1996*b*).

In tandem with the education reforms, the government also implemented significant legislative changes to transform the labour market. Key among these were the Labour Relations Act (LRA) of 1995 (Republic of South Africa, 1995), the Basic Conditions of Employment Act (BCEA) (Republic of South Africa, 1997), the Employment Equity Act (EEA) of 1998 (Republic of South Africa, 1998), and the Broad-Based Black Economic Empowerment (B-BBEE) Act of 2003 (Republic of South Africa, 2004). The LRA expanded workers' rights by institutionalizing collective bargaining, enabling union membership, and protecting the right to strike.

The BCEA aimed to create fairer employment conditions by mandating minimum standards, such as annual leave, sick leave, severance pay, and other basic entitlements. Importantly, the BCEA also introduced sector-specific minimum wages for vulnerable workers in industries with low union representation or low average earnings. These industries included domestic work, contract cleaning, private security, wholesale and retail, farming, forestry, taxis, learnerships, hospitality, civil engineering, and children's roles in advertising, creative arts, and cultural endeavors (Millea et al., 2017; Nattrass, 2016). However, it was not until 2019 that a comprehensive national minimum wage law was enacted, signaling the slow yet progressive journey towards addressing economic inequalities.

The impact of these sector-specific minimum wage laws on employment and wages remains an area of extensive study and debate, with findings offering a mixed picture. Some studies report no statistically significant effects on employment rates following the implementation of minimum wage laws. For instance, research by Bhorat et al. (2013); Conradie (2003, 2005); Murray and Van Walbeek (2007) found limited evidence of employment changes across various sectors. Conversely, other studies, such as Bhorat, Kanbur and Stanwix (2014), suggest that minimum wage laws reduced the likelihood of employment, particularly in certain vulnerable industries.

The effect of minimum wage policies on wages is similarly inconclusive. Bhorat et al. (2013) found that minimum wage laws successfully raised real hourly wages in sectors like private security, retail, and domestic work. However, the same study noted wage decreases in the taxi sector and statistically insignificant impacts in forestry. In the agricultural sector, minimum wage regulations introduced notable wage increases for farm workers, as evidenced by Bhorat, Kanbur and Stanwix (2014). These findings highlight the sector-specific nature of minimum wage policies and their varying implications for workers and employers, underscoring the need for nuanced approaches to labour market interventions.

Unemployment levels in South Africa are high with over 40% of the labour force without jobs (see Figure 1.1). High levels of unemployment have significantly impacted the socioeconomic landscape of the country by contributing to elevated levels of income inequality and poverty (Anand et al., 2016). This issue affects not only individuals but also has a widespread effect on communities, leading to a cycle of deprivation that is hard to break (The World Bank, 2018).

Although access to educational opportunities for the formerly disadvantaged non-White population has improved with the introduction of a non-racial education system and compulsory schooling (UNESCO, 2020), educational outcomes remain poor, particularly within the Black African and Coloured communities. These groups often face numerous challenges, including under-resourced schools, inadequate facilities, and a lack of qualified teachers, which hinder academic progress and limit students' future job prospects (Spaull, 2013).

Therefore, South Africa provides a unique context to study the impact of education on the labour market. This thesis addresses three specific objectives. The first objective is to investigate whether education offers a channel for improving employment outcomes among men and women within the Black African, Coloured, and White communities (the Indian/Asian race was excluded due to small sample size). The second is to assess whether exposure to the compulsory schooling reform increases returns to education at different locations of the wage distribution for Black South Africans. The analysis is performed for both men and women. And the third objective is to estimate the impact of education on wages of Black men and women South Africans.

In Chapter 2 I took advantage of the compulsory schooling policy of 1997 and used it as a source of exogenous variation to identify the causal effect of education on employment outcomes. The instrumental variable is measured as the number of years of compulsory schooling an individual was exposed to. Using a two-stage least squares method to estimate the causal effect of education on employment with data from the General Household Surveys for the years 2014-2018, I showed that the instrument is plausible since it has a strong correlation with the variable years of education.

Also, the IV is exogenous since the compulsory education policy was merely intended to raise the educational levels of the underserved population groups, and not designed to improve employment outcomes directly. Thus, as an IV, it affects the years of education, but is unrelated to ability, which is omitted in the analyses due to lack of data. The analysis of employment outcomes provides evidence that education is endogenous, indicating that earlier South African research probably have endogeneity bias.

The findings indicate that an additional year of education raises the likelihood of employment by 7.3 to 28.4 percentage points, varying across survey years. The estimates point to notable differences by gender and race — both in the effect of compulsory schooling on educational attainment and in the impact of education on employment prospects. Disaggregating by gender, the employment gains linked to an additional year of education range from 3.0 to 18.0 percentage points for women and 3.0 to 13.9 percentage points for men. These results underscore the importance of education in enhancing employment prospects in South Africa. Strengthening access to quality education, especially at higher levels, can be an effective policy lever for improving labour market outcomes. By promoting educational attainment, particularly among disadvantaged groups, policymakers can help increase individuals' chances of employment and contribute to addressing South Africa's persistently high unemployment rates.

The analysis of returns to education along the wage distribution in Chapter 3 involved estimating the impact of compulsory schooling on returns to education along the distribution of earnings. I applied the quantile regression and the quantile selection model to control for selection into employment to find the rate of return to education at different locations of the earnings distribution for both Black men and women who were exposed and those who were not exposed to the compulsory schooling reform.

The findings indicate that returns to education vary based on exposure to compulsory schooling, educational level, gender, and across the wage distribution. Returns are higher for women than for men. In addition, women exposed to the compulsory schooling reform tend to earn more than those who were not. A similar pattern is observed among men, but only at certain quantiles of the wage distribution.

As expected, I find that female returns suffer from sample selection bias, and correcting for the selection lowers their returns, depending on the location on the wage distribution. Moreover, the results suggest a skills-biased labour market that increases wage inequality. This is supported by the analysis that is based on levels of education, rather than years of education. I find increasing returns along educational levels, rising from the lower to the higher end of the wage distribution both for females and males. This increasing pattern of returns to education suggests a skills-biased structure to the labour market, which increases wage inequality.

Chapter 4 explores the impact of education on labour market earnings among Black South Africans. Under South Africa's compulsory schooling law, children are required to begin Grade R or Grade 1 in the calendar year they turn six or seven, provided their birthday falls on or before June 30 of that admission year. Leveraging the natural variation in educational attainment induced by this school entry cutoff, I estimate the returns to education using a regression discontinuity design based on data from the General Household Survey (GHS) spanning 2014 to 2018. This study contributes to the limited body of research on the effects of school entry policies within Sub-Saharan Africa.

The analysis detects indications of birth date manipulation around January 1^{st}, likely intended to circumvent school admission regulations. To mitigate bias arising from this manipulation, individuals born on January 1^{st} are excluded from the estimation sample. This exclusion helps to reduce potential distortions in the estimated treatment effects. The findings suggest that starting school at a younger age is associated with fewer completed years of schooling and lower earnings in the labour market. Nevertheless, for those who do enter school early, education exerts a statistically significant positive influence on earnings. These results raise important policy questions regarding the optimal age for school entry and underscore the need for targeted support to younger entrants who may experience academic disadvantages.

BIBLIOGRAPHY

Aakvik, A., Salvanes, K. G. and Vaage, K. (2010), 'Measuring heterogeneity in the returns to education using an education reform', *European Economic Review* **54**(4), 483–500. https://doi.org/10.1016/j.euroecorev.2009.09.001.

Abekah-Nkrumah, G., Asuming, P. O. and Yusif, H. (2019), 'Duration of high school education and youth labour market outcomes: Evidence from a policy experiment in Ghana', *Journal of International Development* **31**(7), 617–631. https://doi.org/10.1002/jid.3429.

Acemoglu, D. and Angrist, J. (2000), 'How large are human-capital externalities? evidence from compulsory schooling laws', *NBER Macroeconomics Annual* **15**, 9–59. https://doi.org/10.1086/654403.

Alam, K. and Mamun, S. A. K. (2016), 'The relationship between labour force status and educational attainment: Evidence from a system of simultaneous equations model', *Economic Analysis and Policy* **52**, 55–65. https://doi.org/10.1016/j.eap.2016.07.005.

Anand, R., Kothari, S. and Kumar, N. (2016), *South Africa: Labor market dynamics and inequality*, International Monetary Fund.

Andini, C. (2010), 'Within-groups wage inequality and schooling: further evidence for Portugal', *Applied Economics* **42**(28), 3685–3691. https://doi.org/10.1080/00036840802314564.

Angrist, J. D. and Keueger, A. B. (1991), 'Does compulsory school attendance affect schooling and earnings?', *The Quarterly Journal of Economics* **106**(4), 979–1014.
https://doi.org/10.2307/2937954.

Angrist, J. D. and Krueger, A. B. (1992), 'The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples', *Journal of the American statistical Association* **87**(418), 328–336.
https://doi.org/10.2307/2290263.

Arellano, M. and Bonhomme, S. (2017), 'Quantile selection models with an application to understanding changes in wage inequality', *Econometrica* **85**(1), 1–28.
https://doi.org/10.3982/ECTA14030.

Arrow, K. J. et al. (1973), 'Higher education as a filter', *Journal of Public Economics* **2**(3), 193–216.
https://doi.org/10.1016/0047-2727(73)90013-3.

Ashenfelter, O. and Krueger, A. (1994), 'Estimates of the economic return to schooling from a new sample of twins', *The American Economic Review* pp. 1157–1173.
https://www.jstor.org/stable/2117766.

Aydemir, A. and Kirdar, M. G. (2017), 'Low wage returns to schooling in a developing country: Evidence from a major policy reform in Turkey', *Oxford Bulletin of Economics and Statistics* **79**(6), 1046–1086.
https://doi.org/10.1111/obes.12174.

Baah-Boateng, W. (2013), 'Determinants of unemployment in Ghana', *African Development Review* **25**(4), 385–399.
https://doi.org/10.1111/1467-8268.12037.

Baah-Boateng, W. (2015), 'Unemployment in Ghana: a cross sectional analysis from demand and supply perspectives', *African Journal of Economic and Management Studies* .
https://doi.org/10.1108/AJEMS-11-2014-0089.

Baigrie, N. and Eyal, K. (2014), 'An evaluation of the determinants and implications of panel attrition in the national income dynamics survey (2008-2010)', *South African Journal of Economics* **82**(1), 39–65.
https://doi.org/10.1111/saje.12024.

Baker and McKenzie (2023), 'Guidebook to broad-based black economic empowerment'.

Baldry, K. (2016), 'Graduate unemployment in South Africa: social inequality reproduced', *Journal of Education and Work* **29**(7), 788–812.
https://doi.org/10.1080/13639080.2015.1066928.

Balestra, S. and Backes-Gellner, U. (2017), 'Heterogeneous returns to education over the wage distribution: Who profits the most?', *Labour Economics* **44**, 89–105.
https://doi.org/10.1016/j.labeco.2017.01.001.

Balestra, S., Eugster, B. and Liebert, H. (2020), 'Summer-born struggle: The effect of school starting age on health, education, and work', *Health economics* **29**(5), 591–607.
https://doi.org/10.1002/hec.4005.

Bardasi, E. and Taylor, M. (2008), 'Marriage and wages: A test of the specialization hypothesis', *Economica* **75**(299), 569–591.
https://doi.org/10.1111/j.1468-0335.2007.00630.x.

Barreca, A. I., Guldi, M., Lindo, J. M. and Waddell, G. R. (2011), 'Saving babies? revisiting the effect of very low birth weight classification', *The Quarterly Journal of Economics* **126**(4), 2117–2123.
https://doi.org/10.1093/qje/qjr042.

Beck, K. C. (2025), 'The later the better? a novel approach to estimating the effect of school starting age on adhd and academic skills', *European Sociological Review* p. jcaf004.
https://doi.org/10.1093/esr/jcaf004.

Becker, G. S. (1962), 'Investment in human capital: A theoretical analysis', *Journal of Political Economy* **70**(5, Part 2), 9–49.

https://doi.org/10.1086/258724.

Berhe, M. W. (2021), 'Empirical analysis of urban youth unemployment in Ethiopia', *African Development Review* .
https://doi.org/10.1111/1467-8268.12514.

Bhorat, H. (2000), *Wage premia and wage differentials in the South African labour market*, University of Cape Town.
https://open.uct.ac.za/items/56f7acb6-08a4-4a36-b29c-b6bcd1ea7f92.

Bhorat, H. and Goga, S. (2012), 'The gender wage gap in the post-apartheid South African labour market'.
http://dx.doi.org/10.2139/ssrn.2184258.

Bhorat, H., Goga, S. and Stanwix, B. (2014), Skills-biased labour demand and the pursuit of inclusive growth in South Africa, Technical report, WIDER working paper.
https://doi.org/10.35188/UNU-WIDER/2014/851-3.

Bhorat, H., Kanbur, R. and Mayet, N. (2013), 'The impact of sectoral minimum wage laws on employment, wages, and hours of work in South Africa', *IZA Journal of Labor & Development* **2**(1), 1–27.
http://www.izajold.com/content/2/1/1.

Bhorat, H., Kanbur, R. and Stanwix, B. (2014), 'Estimating the impact of minimum wages on employment, wages, and non-wage benefits: the case of agriculture in South Africa', *American Journal of Agricultural Economics* **96**(5), 1402–1419.
https://doi.org/10.1093/ajae/aau049.

Bhorat, H. and Kimani, E. (2017), 'The role of post-school education and training institutions in predicting labour market outcomes', *HSRC LIMP Report* **23**.

Bhorat, H., Stanwix, B. et al. (2021), 'The impact of the national minimum wage in South Africa: Early quantitative evidence'.
https://coilink.org/20.500.12592/1w7g30.

Biyase, M. and Zwane, T. (2015), 'Does education pay in South Africa? estimating returns to education using two stage least squares approach', *The International Business & Economics Research Journal (Online)* **14**(6), 807.
https://core.ac.uk/download/pdf/268108007.pdf.

Blau, F. D., Kahn, L. M., Boboshko, N. and Comey, M. (2024), 'The impact of selection into the labor force on the gender wage gap', *Journal of Labor Economics* **42**(4), 1093–1133.
https://doi.org/10.1086/725032.

Boahen, E. A., Opoku, K. and Schotte, S. (2021), 'Duration of pre-university education and labour market outcomes: Evidence from a quasi-experiment in Ghana', *Journal of International Development* **33**(1), 208–232.
https://doi.org/10.1002/jid.3519.

Boockmann, B. and Steiner, V. (2006), 'Cohort effects and the returns to education in West Germany', *Applied Economics* **38**(10), 1135–1152.
https://doi.org/10.1080/00036840500439168.

Bosch, A. and Koch, S. F. (2020), 'The South African financial cycle and its relation to household deleveraging', *South African Journal of Economics* **88**(2), 145–173.
https://doi.org/10.1111/saje.12245.

Bound, J., Jaeger, D. A. and Baker, R. M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American statistical association* **90**(430), 443–450.
https://doi.org/10.1080/01621459.1995.10476536.

Branson, N., Ardington, C., Lam, D. and Leibbrandt, M. (2013), 'Changes in education, employment and earnings in South Africa: A cohort analysis'.
http://hdl.handle.net/11090/621.

Branson, N., Garlick, J., Lam, D. and Leibbrandt, M. (2012), 'Education and inequality: The South African case'.
http://hdl.handle.net/11090/168.

Branson, N., Hofmeyr, C. and Lam, D. (2014), 'Progress through school and the determinants of school dropout in South Africa', *Development Southern Africa* **31**(1), 106–126.
https://doi.org/10.1080/0376835X.2013.853610.

Branson, N. and Leibbrandt, M. (2013*a*), 'Education quality and labour market outcomes in South Africa'.
https://doi.org/10.1787/5k4c0vvbvv0q-en.

Branson, N. and Leibbrandt, M. (2013*b*), 'Educational attainment and labour market outcomes in South Africa, 1994-2010'.
https://doi.org/10.1787/5k4c0vvbvv0q-en.

Brunello, G., Fort, M. and Weber, G. (2009), 'Changes in compulsory schooling, education and the distribution of wages in Europe', *The Economic Journal* pp. 516–539.
https://doi.org/10.1111/j.1468-0297.2008.02244.x.

Buchinsky, M. (1994), 'Changes in the US wage structure 1963-1987: Application of quantile regression', *Econometrica* pp. 405–458.
https://doi.org/10.2307/2951618.

Buchinsky, M. (1998), 'The dynamics of changes in the female wage distribution in the USA: a quantile regression approach', *Journal of Applied Econometrics* **13**(1), 1–30.
https://doi.org/10.1002/(SICI)1099-1255(199801/02)13:1<1::AID-JAE474>3.0.CO;2-A.

Buchinsky, M. (2001), 'Quantile regression with sample selection: Estimating women's return to education in the US', *Empirical Economics* **26**(1), 87–113.

Burger, R., Jafta, R. et al. (2010), 'Affirmative action in South Africa: an empirical assessment of the impact on labour market outcomes', *CRISE (Centre for Research on Inequality, Human Security and Ethnicity) Working Paper* **76**, 09–36.

Burger, R., Van der Berg, S. and Von Fintel, D. (2015), 'The unintended consequences of education policies on South African participation and unemployment', *South African Journal of Economics* **83**(1), 74–100.
https://doi.org/10.1111/saje.12049.

Burger, R., Von Fintel, D. et al. (2009), 'Determining the causes of the rising South African unemployment rate: An age, period and generational analysis', *Economic Research Southern Africa (ERSA) Working Papers* **158**.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2017), 'rdrobust: Software for regression-discontinuity designs', *The Stata Journal* **17**(2), 372–404.
https://doi.org/10.1177/1536867X1701700208.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014*a*), 'Robust data-driven inference in the regression-discontinuity design', *The Stata Journal* **14**(4), 909–946.
https://doi.org/10.1177/1536867X1401400413.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014*b*), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.
https://doi.org/10.3982/ECTA11757.

Card, D. (2001), 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica* **69**(5), 1127–1160.
https://doi.org/10.1111/1468-0262.00237.

Casale, D. (2004), 'What has the feminisation of the labour market 'bought'women in South Africa? Trends in labour force participation, employment and earnings, 1995–2001', *Journal of Interdisciplinary Economics* **15**(3-4), 251–275.
https://doi.org/10.1177/02601079X04001500302.

Case, A. and Deaton, A. (1999), 'School inputs and educational outcomes in South Africa', *The Quarterly Journal of Economics* **114**(3), 1047–1084.
https://doi.org/10.1162/003355399556124.

Case, A. and Yogo, M. (1999), 'Does school quality matter? returns to education and the characteristics of schools in South Africa'.
https://doi.org//10.3386/w7399.

Castel, V., Phiri, M. and Stampini, M. (2010), 'Education and employment in Malawi', *African Development Bank Working Paper* .

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2020), 'A practical introduction to regression discontinuity designs: Extensions', *Cambridge Elements: Quantitative and Computational Methods for Social Science* .
https://doi.org/10.1017/9781009441896.

Cattaneo, M. D., Jansson, M. and Ma, X. (2020), 'Simple local polynomial density estimators', *Journal of the American Statistical Association* **115**(531), 1449–1455.
https://doi.org/10.1080/01621459.2019.1635480.

Chiswick, B. R. (2024), 'Estimating returns to schooling and experience: a history of thought', *Education Economics* **32**(5), 566–580.
https://doi.org/10.1080/09645292.2024.2326015.

Chuanchuan, Z. and Jingwen, W. (2021), 'Gender roles and women's labor market outcomes', *China Economic Quarterly International* **1**(2), 97–108.
https://doi.org/10.1016/j.ceqi.2021.04.002.

Clark, D. (2023), 'School quality and the return to schooling in Britain: New evidence from a large-scale compulsory schooling reform', *Journal of Public Economics* **223**, 104902.
https://doi.org/10.1093/ajae/aau049.

Clark, D. and Martorell, P. (2014), 'The signaling value of a high school diploma', *Journal of Political Economy* **122**(2), 282–318.
https://doi.org/10.1086/675238.

Clay, K., Lingwall, J. and Stephens Jr, M. (2021), 'Laws, educational outcomes, and returns to schooling evidence from the first wave of US state compulsory attendance laws', *Labour Economics* **68**, 101935.
https://doi.org/10.1016/j.labeco.2020.101935.

Conradie, B. (2003), 'Labour, wages and minimum wage compliance in the Breërivier valley six months after the introduction of minimum wages', *Centre for Social Science Research Working Paper No. 51* .
https://hdl.handle.net/11427/19725.

Conradie, B. (2005), 'Wages and wage elasticities for wine and table grapes in South Africa', *Agrekon* **44**(1), 138–156.
https://doi.org/10.1080/03031853.2005.9523706.

Dai, F., Cai, F. and Zhu, Y. (2022), 'Returns to higher education in China–evidence from the 1999 higher education expansion using a fuzzy regression discontinuity', *Applied Economics Letters* **29**(6), 489–494.
https://doi.org/10.1080/13504851.2020.1871465.

de New, S. C., Schurer, S. and Sulzmaier, D. (2021), 'Gender differences in the lifecycle benefits of compulsory schooling policies', *European Economic Review* **140**, 103910.
https://doi.org/10.1016/j.euroecorev.2021.103910.

Department of Employment and Labour (2023), 'Commission for Employment Equity: Annual Report 2022 - 2023'.

Depken, C., Chiseni, C. and Ita, E. (2019), 'Returns to education in South Africa: Evidence from the National Income Dynamics Study', *Zagreb International Review of Economics & Business* **22**(1), 1–12.
https://doi.org/10.2478/zireb-2019-0009.

Devereux, P. J. and Hart, R. A. (2010), 'Forced to be rich? returns to compulsory schooling in Britain', *The Economic Journal* **120**(549), 1345–1364.
https://doi.org/10.1111/j.1468-0297.2010.02365.x.

Diawara, B., Mughal, M. et al. (2011), 'Education and employment in Sub-Saharan Africa: Some evidence from Senegal', *Theoretical and Applied Economics* **5**(5), 59.

Dobkin, C. and Ferreira, F. (2010), 'Do school entry laws affect educational attainment and labor market outcomes?', *Economics of education review* **29**(1), 40–54. https://doi.org/10.1016/j.econedurev.2009.04.003.

Dolado, J. J., García-Peñalosa, C. and Tarasonis, L. (2020), 'The changing nature of gender selection into employment over the great recession', *Economic Policy* **35**(104), 635–677. https://doi.org/10.1093/epolic/eiaa025.

Dolton, P. and Sandi, M. (2017), 'Returning to returns: Revisiting the British education evidence', *Labour Economics* **48**, 87–104. https://doi.org/10.1016/j.labeco.2017.07.001.

Dougherty, C. (2005), 'Why are the returns to schooling higher for women than for men?', *Journal of Human Resources* **40**(4), 969–988. https://doi.org/10.3368/jhr.XL.4.969.

Draissi, Z., Rong, Y. and Ebrahim Suliman, M. A. (2023), 'Employability and education mismatch in the Moroccan urban workplace', *SAGE Open* **13**(4), 21582440231217890. https://doi.org/10.1177/21582440231217890.

Drucker, L. F., Horn, D. and Jakubowski, M. (2022), 'The labour market effects of the Polish educational reform of 1999', *Journal for Labour Market Research* **56**(1), 1–21. https://doi.org/10.1186/s12651-022-00317-4.

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.

Elass, K. (2024), 'Male and female selection effects on gender wage gaps in three countries', *Labour Economics* **87**, 102506. https://doi.org/10.1016/j.labeco.2024.102506.

Espi, G., Francis, D. and Valodia, I. (2019), 'Gender inequality in the South African labour market: Insights from the employment equity act data', *Agenda* **33**(4), 44–61.
https://doi.org/10.1080/10130950.2019.1674675.

Fan, E., Meng, X., Wei, Z. and Zhao, G. (2018), 'Rates of return to four-year university education: An application of regression discontinuity design', *The Scandinavian Journal of Economics* **120**(4), 1011–1042.
https://doi.org/10.1111/sjoe.12251.

Fasih, T., Kingdon, G. G., Patrinos, H. A., Sakellariou, C. and Söderbom, M. (2012), 'Heterogeneous returns to education in the labor market'.
https://ssrn.com/abstract=2129612.

Festus, L., Kasongo, A., Moses, M. and Yu, D. (2016), 'The South African labour market, 1995–2015', *Development Southern Africa* **33**(5), 579–599.
https://doi.org/10.1080/0376835X.2016.1203759.

Finn, A. (2015), 'A National Minimum Wage in the Context of the South African Labour Market', Technical report, Southern Africa Labour and Development Research Unit Working Paper Number 153, University of Cape Town.

Finn, A. and Leibbrandt, M. (2018), The evolution and determination of earnings inequality in post-apartheid South Africa, Technical report, WIDER Working Paper.
https://hdl.handle.net/10419/190132.

Francis, D. and Webster, E. (2019), 'Poverty and inequality in South Africa: critical reflections', *Development Southern Africa* **36**(6), 788–802.
https://doi.org/10.1080/0376835X.2019.1666703.

Fredriksson, P. and Öckert, B. (2014), 'Life-cycle effects of age at school start', *The Economic Journal* **124**(579), 977–1004.
https://doi.org/10.1111/ecoj.12047.

Fryer, D. and Vencatachellum, D. (2005), 'Returns to education in South Africa: Evidence from the Machibisa township', *African Development Review* **17**(3), 513–535. https://doi.org//10.1111/j.1017-6772.2006.00126.x.

Gelman, A. and Imbens, G. (2019), 'Why high-order polynomials should not be used in regression discontinuity designs', *Journal of Business & Economic Statistics* **37**(3), 447–456. https://doi.org/10.1080/07350015.2017.1366909.

Giotis, G., Gogas, T. and Gouda, K. (2025), 'Education and employment', *Encyclopedia* **5**(2), 85. https://doi.org/10.3390/.

Girma, S. and Kedir, A. (2005), 'Heterogeneity in returns to schooling: Econometric evidence from Ethiopia', *The Journal of Development Studies* **41**(8), 1405–1416. https://doi.org/10.1080/00220380500187026.

González, F. A. I. and Dip, J. A. (2024), 'School entry cutoff and the timing of births: evidence from Argentina', *Education Economics* **32**(1), 47–58. https://doi.org/10.1080/09645292.2023.2173148.

Görlitz, K., Heß, P. and Tamm, M. (2025), 'Should states allow early school enrollment? an analysis of individuals' long-term labor market effects', *Empirical Economics* pp. 1–29. https://doi.org/10.1007/s00181-024-02702-z.

Gradín, C. (2019), 'Occupational segregation by race in South Africa after apartheid', *Review of Development Economics* **23**(2), 553–576. https://doi.org/10.1111/rode.12551.

Granados, P. G., Wrohlich, K. et al. (2020), *Selection into Employment and the Gender Wage Gap across the Distribution and over Time*, Deutsches Institut für Wirtschaftsforschung (DIW). https://hdl.handle.net/10419/215255.

Grenet, J. (2013), 'Is extending compulsory schooling alone enough to raise earnings? evidence from French and British compulsory schooling laws', *The Scandinavian Journal of Economics* **115**(1), 176–210.

https://doi.org/10.1111/j.1467-9442.2012.01739.x.

Gronau, R. (1974), 'Wage comparisons–a selectivity bias', *Journal of Political Economy* **82**(6), 1119–1143.
https://doi.org/10.1086/260267.

Guo, C., Wang, X. and Meng, C. (2023), 'Does the early bird catch the worm? evidence and interpretation on the long-term impact of school entry age in China', *China Economic Review* **77**, 101900.
https://doi.org/10.1016/j.chieco.2022.101900.

Hall, C. (2016), 'Does more general education reduce the risk of future unemployment? evidence from an expansion of vocational upper secondary education', *Economics of Education Review* **52**, 251–271.
https://doi.org/10.1016/j.econedurev.2016.03.005.

Harmon, C., Oosterbeek, H. and Walker, I. (2000), *The returns to education: a review of evidence, issues and deficiencies in the literature*, Centre for the Economics of Education, London School of Economics and Political Science.

Harmon, C., Oosterbeek, H. and Walker, I. (2003), 'The returns to education: Microeconomics', *Journal of economic surveys* **17**(2), 115–156.

Harmon, C. P. (2017), 'How effective is compulsory schooling as a policy instrument?', *IZA World of Labor* .
https://doi.org//10.15185/izawol.348.

Harmon, C. and Walker, I. (1995), 'Estimates of the economic return to schooling for the United Kingdom', *The American Economic Review* **85**(5), 1278–1286.
https://www.jstor.org/stable/2950988.

Hartog, J., Pereira, P. T. and Vieira, J. A. (2001), 'Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators', *Applied Economics* **33**(8), 1021–1037.

https://doi.org/10.1080/00036840122679.

Heckman, J. (1974), 'Shadow prices, market wages, and labor supply', *Econometrica* pp. 679–694. https://doi.org/10.2307/1913937.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica* pp. 153–161. https://doi.org/10.2307/1912352.

Hertz, T. (2003), 'Upward bias in the estimated returns to education: Evidence from South Africa', *American Economic Review* **93**(4), 1354–1368. https://doi.org//10.1257/000282803769206331.

Hinks, T. (2002), 'Gender wage differentials and discrimination in the new South Africa', *Applied Economics* **34**(16), 2043–2052. https://doi.org/10.1080/00036840210124991.

Horwitz, F. M. and Jain, H. (2011), 'An assessment of employment equity and broad based black economic empowerment developments in South Africa', *Equality, Diversity and Inclusion: An International Journal* **30**(4), 297–317. https://doi.org/10.1108/02610151111135750.

Huang, C., Zhang, S. and Zhao, Q. (2020), 'The early bird catches the worm? school entry cutoff and the timing of births', *Journal of Development Economics* **143**, 102386. https://doi.org/10.1016/j.jdeveco.2019.102386.

Huber, M. and Melly, B. (2015), 'A test of the conditional independence assumption in sample selection models', *Journal of Applied Econometrics* **30**(7), 1144–1168. https://doi.org/10.1002/jae.2431.

Huntington-Klein, N. (2021), 'Human capital versus signaling is empirically unresolvable', *Empirical Economics* **60**(5), 2499–2531. https://doi.org/10.1007/s00181-020-01837-z.

ILO (2024), *World employment and social outlook; Trends 2024*, ILO, Geneva. https://doi.org/10.54394/HQAE1085.

Kawaguchi, D. (2011), 'Actual age at school entry, educational outcomes, and earnings', *Journal of the Japanese and International Economies* **25**(2), 64–80.
https://doi.org/10.1016/j.jjie.2009.02.002.

Kaymak, B. (2025), 'Quantifying the signaling role of education', *Federal Reserve Bank of Cleveland Working Paper Series Working Paper* .
https://doi.org/10.26509/frbc-wp-202502.

Kercheval, J., Markowitz, D., Monson, K., Read, B. and Brown, D. (2013), 'Women in leadership—research on barriers to employment and decent work for women. literature review', *International Labour Organization (ILO)* .

Keswell, M. and Poswell, L. (2002), 'How important is education for getting ahead in South Africa?'.
http://hdl.handle.net/11427/20157.

Keswell, M. and Poswell, L. (2004), 'Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence', *South African Journal of Economics* **72**(4), 834–860.
https://doi.org//10.1111/j.1813-6982.2004.tb00136.x.

Kingdon, G. G. and Knight, J. (2004), 'Race and the incidence of unemployment in South Africa', *Review of Development Economics* **8**(2), 198–222.
https://doi.org/10.1111/j.1467-9361.2004.00228.x.

Klasen, S. and Minasyan, A. (2021), 'Affirmative action and intersectionality at the top: Evidence from South Africa', *Industrial Relations* **60**(1), 3–35.
https://doi.org/10.1111/irel.12272.

Koenker, R. and Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica* pp. 33–50.
https://doi.org/10.2307/1913643.

Kollamparambil, U. and Razak, A. (2016), 'Trends in gender wage gap and discrimination in South Africa: A comparative analysis across races', *Indian Journal of Human Development* **10**(1), 49–63.
https://doi.org/10.1177/0973703016636446.

Korwatanasakul, U. (2023), 'Returns to schooling in Thailand: Evidence from the 1978 compulsory schooling law', *The Developing Economies* **61**(1), 3–35.
https://doi.org/10.1111/deve.12341.

Kuepié, M. and Nordman, C. J. (2016), 'Where does education pay off in Sub-Saharan Africa? Evidence from two cities of the Republic of Congo', *Oxford Development Studies* **44**(1), 1–27.
https://doi.org/10.1080/13600818.2015.1110568.

Lam, D. (2016), 'The impact of race on earnings and human capital in Brazil, South Africa and the United States', *Anais* pp. 1–26.

Lam, D., Ardington, C. and Leibbrandt, M. (2011), 'Schooling as a lottery: Racial differences in school advancement in urban South Africa', *Journal of Development Economics* **95**(2), 121–136.
https://doi.org/10.1016/j.jdeveco.2010.05.005.

Lam, D., Leibbrandt, M. and Mlatsheni, C. (2009), Education and youth unemployment in South Africa, *in* 'Labor Markets and Economic Development', Routledge, pp. 90–109.

Landman, M. S., O'Clery, N. et al. (2020), The impact of the employment equity act on female inter-industry labour mobility and the gender wage gap in South Africa, Technical report, World Institute for Development Economic Research (UNU-WIDER).
https://hdl.handle.net/10419/229276.

Larsen, E. R. and Solli, I. F. (2017), 'Born to run behind? persisting birth month effects on earnings', *Labour Economics* **46**, 200–210.
http://dx.doi.org/10.1016/j.labeco.2016.10.005.

Lewis, B. D. and Nguyen, H. T. (2020), 'Assessing the causal impact of compulsory schooling policy in Indonesia', *International Journal of Educational Research* **104**, 101693.

https://doi.org/10.1016/j.ijer.2020.101693.

Li, X., Lou, X. and Zhang, J. (2022), 'Does the early bird catch the worm? the effect of school starting age on educational attainment and labor market outcomes: Evidence from Chinese urban twins data', *Journal of Comparative Economics* **50**(3), 832–848. https://doi.org/10.1016/j.jce.2022.04.004.

Liu, S. and Marois, G. (2023), 'The effect of motherhood on the labour force participation of married women in China', *Asian Population Studies* pp. 1–17. https://doi.org/10.1080/17441730.2023.2193518.

Liwiński, J. (2020), 'The impact of compulsory education on employment and wages in a transition economy', *Eastern European Economics* **58**(2), 137–173. https://doi.org/10.1080/00128775.2019.1689144.

Maasoumi, E. and Wang, L. (2019), 'The gender gap between earnings distributions', *Journal of Political Economy* **127**(5), 2438–2504. https://doi.org/10.1086/701788.

Marcenaro-Gutierrez, O. D. and Lopez-Agudo, L. A. (2021), 'Too late or too soon for school? The impact of school entry age', *Journal of research on educational effectiveness* **14**(2), 309–352. https://doi.org/10.1080/19345747.2020.1849479.

Markussen, S. and Strøm, M. (2022), 'Children and labor market outcomes: separating the effects of the first three children', *Journal of Population Economics* **35**(1), 135–167. https://doi.org/10.1007/s00148-020-00807-0.

Martins, P. S. and Pereira, P. T. (2004), 'Does education reduce wage inequality? quantile regression evidence from 16 countries', *Labour Economics* **11**(3), 355–371. https://doi.org/10.1016/j.labeco.2003.05.003.

Matta, R., Ribas, R. P., Sampaio, B. and Sampaio, G. R. (2016), 'The effect of age at school entry on college admission and earnings: a regression-discontinuity approach', *IZA Journal of Labor Economics* **5**, 1–25.

https://doi.org/10.1186/s40172-016-0049-5.

McCrary, J. (2008), 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of Econometrics* **142**(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005.

Millea, M. J., Rezek, J. P., Shoup, B. and Pitts, J. (2017), 'Minimum wages in a segmented labor market: Evidence from South Africa', *Journal of Labor Research* **38**, 335–359. https://doi.org/10.1007/s12122-017-9241-z.

Mincer, J. (1958), 'Investment in human capital and personal income distribution', *Journal of Political Economy* **66**(4), 281–302. https://doi.org/10.1086/258055.

Mincer, J. (1974), 'Schooling, experience, and earnings. Human Behavior & Social Institutions No. 2.'.

Mishra, V. and Smyth, R. (2015), 'Estimating returns to schooling in urban China using conventional and heteroskedasticity-based instruments', *Economic Modelling* **47**, 166–173. https://doi.org/10.1016/j.econmod.2015.02.002.

Mitra, A. (2019), 'Returns to education in India: Capturing the heterogeneity', *Asia & the Pacific Policy Studies* **6**(2), 151–169. https://doi.org/10.1002/app5.271.

Mlachila, M. M. and Moeletsi, T. (2019), 'Struggling to make the grade: A review of the causes and consequences of the weak outcomes of South Africa's education system', *IMF Working Papers* .

Moll, P. G. (1992), 'Quality of education and the rise in returns to schooling in South Africa, 1975–1985', *Economics of Education Review* **11**(1), 1–10. https://doi.org/10.1016/0272-7757(92)90016-V.

Moll, P. G. (1998), 'Primary schooling, cognitive skills and wages in South Africa', *Economica* **65**(258), 263–284. https://doi.org/10.1111/1468-0335.00127.

Mosomi, J. (2019*a*), 'Distributional changes in the gender wage gap in the post-apartheid South African labour market'.
https://hdl.handle.net/10419/211247.

Mosomi, J. (2019*b*), 'An empirical analysis of trends in female labour force participation and the gender wage gap in South Africa', *Agenda* **33**(4), 29–43.
https://doi.org/10.1080/10130950.2019.1656090.

Mosomi, J. W. et al. (2020), 'The labor market in South Africa, 2000–2017', *IZA World of Labor* .
https://doi.org/10.15185/izawol.475.

Msila, V. (2007), 'From apartheid education to the revised national curriculum statement: Pedagogy for identity formation and nation building in South Africa', *Nordic Journal of African Studies* **16**(2).
https://doi.org/10.53228/njas.v16i2.63.

Mtapuri, O. and Tinarwo, P. (2021), 'From apartheid to democracy', *Southern African Journal of Demography* **21**(1), 104–133.
https://doi.org/10.2307/27125725.

Muñoz, E. and Siravegna, M. (2021), 'Implementing quantile selection models in stata', *The Stata Journal* **21**(4), 952–971.
https://doi.org/10.1177/1536867X211063148.

Murray, J. and Van Walbeek, C. (2007), 'Impact of the sectoral determination for farm workers on the South African sugar industry: Case study of the KwaZulu-Natal North and South Coasts', *Agrekon* **46**(1), 94–112.
https://doi.org/10.1080/03031853.2007.9523763.

Mwabu, G. and Schultz, T. P. (1996), 'Education returns across quantiles of the wage function: alternative explanations for returns to education by race in South Africa', *The American Economic Review* **86**(2), 335–339.
https://www.jstor.org/stable/2118148.

Mwabu, G. and Schultz, T. P. (2000), 'Wage premiums for education and location of South African workers, by gender and race', *Economic Development and Cultural Change* **48**(2), 307–334. https://doi.org/10.1086/452460.

Nattrass, N. (2016), 'Setting minimum wages in South Africa', *International Labour Organization* **2**, 704.

Naude, W. and Serumaga-Zake, P. (2001), 'An analysis of the determinants of labour force participation and unemployment in South Africa's North-West province', *Development Southern Africa* **18**(3), 261–278. https://doi.org/10.1080/03768350120041929.

Nikolov, P., Jimi, N. and Chang, J. (2020), 'The importance of cognitive domains and the returns to schooling in South Africa: Evidence from two labor surveys', *Labour Economics* **65**, 101849. https://doi.org/10.1016/j.labeco.2020.101849.

Ntuli, M. (2009), Exploring gender wage "discrimination" in South Africa, 1995–2004: a quantile regression approach, *in* 'Labor Markets and Economic Development', Routledge, pp. 498–525.

Ntuli, M. and Wittenberg, M. (2013), 'Determinants of black women's labour force participation in post-apartheid South Africa', *Journal of African Economies* **22**(3), 347–374. https://doi.org/10.1093/jae/ejs039.

Oosterbeek, H., ter Meulen, S. and van Der Klaauw, B. (2021), 'Long-term effects of school-starting-age rules', *Economics of Education Review* **84**, 102144. https://doi.org/10.1016/j.econedurev.2021.102144.

Oreopoulos, P. (2006*a*), 'The compelling effects of compulsory schooling: Evidence from Canada', *Canadian Journal of Economics / Revue canadienne d'économique* **39**(1), 22–52. https://doi.org/10.1111/j.0008-4085.2006.00337.x.

Oreopoulos, P. (2006*b*), 'Estimating average and local average treatment effects of education when compulsory schooling laws really matter', *American Economic Review* **96**(1), 152–175. https://doi.org//10.1257/000282806776157641.

Oreopoulos, P. and Salvanes, K. G. (2011), 'Priceless: The nonpecuniary benefits of schooling', *Journal of Economic Perspectives* **25**(1), 159–184.
https://doi.org/10.1257/jep.25.1.159.

Oyelere, R. U. (2010), 'Africa's education enigma? The Nigerian story', *Journal of Development Economics* **91**(1), 128–139.
https://doi.org/10.1016/j.jdeveco.2009.04.001.

Oyenubi, A. and Mosomi, J. (2024), 'Utility of inequality sensitive measures of the gender wage gap: Evidence from South Africa', *Economic Analysis and Policy* **81**, 576–590.
https://doi.org/10.1016/j.eap.2023.12.017.

Patrinos, H. A. (2016), 'Estimating the return to schooling using the mincer equation', *IZA World of Labor* .

Patrinos, H. A., Psacharopoulos, G. and Tansel, A. (2021), 'Private and social returns to investment in education: the case of Turkey with alternative methods', *Applied Economics* **53**(14), 1638–1658.
https://doi.org/10.1080/00036846.2020.1841086.

Patrinos, H. A., Ridao-Cano, C. and Sakellariou, C. (2006), *Estimating the returns to education: accounting for heterogeneity in ability*, The World Bank.

Patrinos, H. A. and Sakellariou, C. (2006), 'Economic volatility and returns to education in Venezuela: 1992–2002', *Applied Economics* **38**(17), 1991–2005.
https://doi.org/10.1080/00036840500427338.

Pehkonen, J., Viinikainen, J., Böckerman, P., Pulkki-Råback, L., Keltikangas-Järvinen, L. and Raitakari, O. (2015), 'Relative age at school entry, school performance and long-term labour market outcomes', *Applied Economics Letters* **22**(16), 1345–1348.
http://dx.doi.org/10.1080/13504851.2015.1031864.

Peña, P. A. (2017), 'Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood', *Economics of Education Review* **56**, 152–176.

https://doi.org/10.1016/j.econedurev.2016.12.001.

Piek, M., von Fintel, D. and Kirsten, J. (2023), 'The impact of agricultural minimum wages on worker flows in South Africa', *South African Journal of Economics* . https://doi.org/10.1111/saje.12357.

Posel, D. (2001), 'How do households work? migration, the household and remittance behaviour in South Africa', *Social Dynamics* **27**(1), 165–189. https://doi.org/10.1080/02533950108458709.

Posel, D. (2010), 'Households and labour migration in post-apartheid South Africa', *Studies in Economics and Econometrics* **34**(3), 129–141. https://hdl.handle.net/10520/EJC21503.

Posel, D. (2025), 'Who cares? gender differences in social reproduction and well-being in South Africa', *Gender, Work & Organization* . https://doi.org/10.1111/gwao.13249.

Posel, D. and Bruce-Brand, J. (2021), "Only a housewife?'Subjective well-being and homemaking in South Africa', *Journal of Happiness Studies* **22**(1), 323–342. https://doi.org/10.1007/s10902-020-00232-w.

Prieto-Rodriguez, J., Barros, C. P. and Vieira, J. A. (2008), 'What a quantile approach can tell us about returns to education in Europe', *Education Economics* **16**(4), 391–410. https://doi.org/10.1080/09645290801977033.

Psacharopoulos, G. (2024), 'Returns to education: a brief history and an assessment', *Education Economics* **32**(5), 561–565. https://doi.org/10.1080/09645292.2024.2370119.

Psacharopoulos, G. and Patrinos, H. A. (2018), 'Returns to investment in education: a decennial review of the global literature', *Education Economics* **26**(5), 445–458. https://doi.org/10.1080/09645292.2018.1484426.

Rahman, M. and Al-Hasan, M. (2018), 'Returns to schooling in Bangladesh revisited', *The Bangladesh Development Studies* **41**(2), 27–42. https://www.jstor.org/stable/27031074.

Ranchhod, V. and Daniels, R. C. (2021), 'Labour market dynamics in South Africa at the onset of the covid-19 pandemic', *South African Journal of Economics* **89**(1), 44–62. https://doi.org/10.1111/saje.12283.

Republic of South Africa (1995), 'Labour relations act, 1995'.

Republic of South Africa (1996*a*), 'Constitution of the Republic of South Africa', *Government Gazette* **378**(17678).

Republic of South Africa (1996*b*), 'South African Schools Act, 1996', *Government Gazette* **377**(1867).

Republic of South Africa (1997), 'Basic conditions of employment act, 1997', *Government Gazette* **390**(75).

Republic of South Africa (1998), 'Employment equity act, 1998', *Government Gazette* **400**(55).

Republic of South Africa (2004), 'Broad-based black economic empowerment act, 2003', *Government Gazette* **463**(46).

Republic of South Africa (2014), 'Employment equity amendment act, 2013', *Government Gazette* **583**(37238).

Republic of South Africa (2024), 'Basic Education Laws Amendment Act, 2024', *Government Gazette* **711**(51258).

Riddell, W. C. and Song, X. (2011), 'The impact of education on unemployment incidence and re-employment success: Evidence from the US labour market', *Labour Economics* **18**(4), 453–463. https://doi.org/10.1016/j.labeco.2011.01.003.

Sakaue, K., Wokadala, J. and Ogawa, K. (2024), 'Updates on private returns to education in Uganda: evidence from universal primary education policy', *Education Economics* pp. 1–17.

https://doi.org/10.1080/09645292.2024.2345708.

Salisbury, T. (2016), 'Education and inequality in South Africa: Returns to schooling in the post-apartheid era', *International Journal of Educational Development* **46**, 43–52. https://doi.org/10.1016/j.ijedudev.2015.07.004.

Schlotter, M., Schwerdt, G. and Woessmann, L. (2011), 'Econometric methods for causal evaluation of education policies and practices: a non-technical guide', *Education Economics* **19**(2), 109–137. https://doi.org/10.1080/09645292.2010.511821.

Schultz, T. W. (1960), 'Capital formation by education', *Journal of Political Economy* **68**(6), 571–583. https://doi.org/10.1086/258393.

Schultz, T. W. (1961), 'Investment in human capital', *The American Economic Review* **51**(1), 1–17. https://www.jstor.org/stable/1818907.

Seekings, J. (2016), 'The quality of South African wage data: accounting for deductions, debt and disclosure', *Centre for Social Science Research Working Paper No. 397* .

Skovron, C. and Titiunik, R. (2015), 'A practical guide to regression discontinuity designs in political science', *American Journal of Political Science* **2015**, 1–36.

Smith, M. (2023), 'Adam Smith on Growth and Economic Development', *History of Economics Review* **86**(1), 2–15. https://doi.org/10.1080/10370196.2023.2243741.

Solarin, S. A. and Yen, Y. Y. (2016), 'A global analysis of the impact of research output on economic growth', *Scientometrics* **108**(2), 855–874. https://doi.org/10.1007/s11192-016-2002-6.

Solli, I. F. (2017), 'Left behind by birth month', *Education Economics* **25**(4), 323–346. https://doi.org/10.1080/09645292.2017.1287881.

Somani, R. (2021), 'The returns to higher education and public employment', *World Development* **144**, 105471.
https://doi.org/10.1016/j.worlddev.2021.105471.

Spaull, N. (2013), 'Poverty & privilege: Primary school inequality in South Africa', *International Journal of Educational Development* **33**(5), 436–447.
https://doi.org/10.1016/j.ijedudev.2012.09.009.

Spaull, N. (2015*a*), Education quality in South Africa and sub-Saharan Africa: An economic approach, PhD thesis, Stellenbosch: Stellenbosch University.

Spaull, N. (2015*b*), 'Schooling in South Africa: How low-quality education becomes a poverty trap', *South African Child Gauge* **12**(1), 34–41.

Spence, M. (1973), 'Job market signaling', *The Quarterly Journal of Economics* **87**(3), 355–374.
https://doi.org/10.1016/B978-0-12-214850-7.50025-5.

Staneva, A. V., Arabsheibani, G. R. and Murphy, P. (2022), Returns to education in four transition countries, *in* 'Critical Perspectives on Economics of Education', Routledge, pp. 182–212.

Statistics South Africa (2016), 'General household survey 2015 [dataset]'.
Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/715.

Statistics South Africa (2017*a*), 'General household survey 2014 [dataset]'.
Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/526.

Statistics South Africa (2017*b*), 'General household survey 2016 [dataset]'.
Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/621.

Statistics South Africa (2018), 'General household survey 2017 [dataset]'.
Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/723.

Statistics South Africa (2019), 'General household survey 2018 [dataset]'.
Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/801.

Statistics South Africa (2023), 'Census 2022 statistical release'. Available from: https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf.

Stefani, P., Biderman, C. et al. (2006), 'Returns to education and wage differentials in Brazil: a quantile approach', *Economics Bulletin* **9**(1), 1–6.

Stefani, P. C. and Biderman, C. (2009), 'The evolution of the returns to education and wage differentials in Brazil: a quantile approach', *Applied Economics* **41**(11), 1453–1460. https://doi.org/10.1080/00036840601019299.

Stiglitz, J. E. (1975), 'The theory of" screening," education, and the distribution of income', *The American Economic Review* **65**(3), 283–300. https://www.jstor.org/stable/1804834.

Tansel, A. and Bodur, F. B. (2012), 'Wage inequality and returns to education in Turkey: A quantile regression analysis', *Review of Development Economics* **16**(1), 107–121. https://doi.org/10.1111/j.1467-9361.2011.00655.x.

The World Bank (2018), 'Overcoming Poverty and Inequality in South Africa: An Assessment of Drivers, Constraints and Opportunities', Technical report, World Bank.

Thobejane, T. (2013), 'History of apartheid education and the problems of reconstruction in South Africa', *Sociology Study* **3**(1), 1–12.

Timæus, I. M., Simelane, S. and Letsoalo, T. (2013), 'Poverty, race, and children's progress at school in South Africa', *The Journal of Development Studies* **49**(2), 270–284. https://doi.org/10.1080/00220388.2012.693168.

Torun, H. (2018), 'Compulsory schooling and early labor market outcomes in a middle-income country', *Journal of Labor Research* **39**(3), 277–305. https://doi.org/10.1007/s12122-018-9264-0.

UNESCO (2020), Global education monitoring report 2020: inclusion and education: all means all, Technical report, UNESCO, Paris, France.

Valdés, M. T. and Requena, M. (2024), 'The effect of the age at school entry on educational attainment and field of study: an analysis using the Spanish census', *Higher Education* **87**(4), 1061–1083.
https://doi.org/10.1007/s10734-023-01053-5.

van der Berg, S., van Wyk, C., Gustafsson, M., Clayton, R., Böhmer, B., Meyer, H., Qvist, J., Mkhombo, T., Euston-Bown, A., Claassen, J. et al. (2025), 'What SA-SAMS and Lurits data tells us about education: New insights from administrative data'.
http://dx.doi.org/10.2139/ssrn.5181087.

van der Berg, S., Wills, G., Selkirk, R., Adams, C. and Van Wyk, C. (2019), The cost of repetition in South Africa, Technical report, Stellenbosch Economic Working Papers: WP13/2019.

Viljoen, D. J. and Dunga, S. H. (2013), 'Determining the factors that influence female unemployment in a South African township', *International Journal of Social Sciences and Humanity Studies* **5**(1), 63–72.

Vu, L. H., Tran, T. Q. and Phung, T. D. (2021), 'Children and female labor market outcomes in Vietnam', *Heliyon* **7**(7).
https://doi.org/10.1016/j.heliyon.2021.e07508.

Wang, L. (2013), 'How does education affect the earnings distribution in urban China?', *Oxford Bulletin of Economics and Statistics* **75**(3), 435–454.
https://doi.org/10.1111/j.1468-0084.2012.00697.x.

Wilcox, V. and Sahni, H. (2023), 'The effects on labor supply of living with older family members needing assistance with activities of daily living', *Journal of Family and Economic Issues* **44**(4), 900–918.
https://doi.org/10.1007/s10834-022-09880-x.

Wittenberg, M. (2017*a*), 'Measurement of earnings: Comparing South African tax and survey data', Technical report, Southern Africa Labour and Development Research Unit, University of Cape Town.

Wittenberg, M. (2017*b*), 'Wages and wage inequality in South Africa 1994–2011: Part 2–inequality measurement and trends', *South African Journal of Economics* **85**(2), 298–318. https://doi.org/10.1111/saje.12147.

Zwane, T. T. (2020), 'The causal effect of education on earnings in urban and rural South Africa: A further update', *Demography and social economy* **39**(1), 79–94.