When the Revolution Hits the University: College Aspirations at the Time of Upheaval

Mahdi Majbouri[*]

Babson College and IZA

Abstract:

A fundamental problem in education policy is how much educational outcomes are affected by government interventions on the supply side of the market. This paper is a response to this policy debate using a natural experiment setting in one of the most under-studied countries, Iran. It measures the effect of Iranian Cultural Revolution during which institutions of higher education were closed and supply of higher education was completely eliminated for 30 months. Using a regression discontinuity design, this paper documents the causal effect of this temporary elimination of supply of higher education on college attainment rates of affected cohorts. The results show that there is a small impact on men's college attainment rate (about 10% or 1.5 percentage points) and no evidence of impact on women's. This corroborates the argument that government interventions have little impact on educational outcomes when demand for education is high. It has important implications for higher education policy in both developing and developed countries. The Cultural Revolution is then used as an instrument to estimate return to college education for men.

JEL Classification Code: I23, I28, C99

Key Words: Higher education, Natural experiment, Government Policy, Institutions

---

[*] Assistant Professor of Economics at Babson College and Research Fellow at IZA. 231 Forest St., Babson Park, MA 02472. Phone: 781-239-5549. Fax: 781-239-5239. mmajbouri@babson.edu

1. Introduction

A principal problem in education policy is how much educational outcomes are affected by government interventions on the supply side of the market. There is a debate on whether expansion of supply increases educational attainment. One view is that because low supply means less access and higher costs, interventions on the supply side are necessary to expand access and increase educational attainments. The opposing view is that if the return to education is high enough, there will be enough demand for education and markets react naturally to increase supply of education without an intervention. Therefore, although interventions in the market to increase access may raise educational attainment, they are not cost-effective (because returns to education is low to begin with.) The debate over this problem is particularly under-researched in higher education.

This paper is a response to this policy debate using a natural experiment setting in one of the most under-studied countries, Iran. It measures the effect of Iranian *Cultural Revolution* during which all institutions of higher education were closed for 30 months. Therefore, the supply of higher education was eliminated for two years and a half. Six birth cohorts who were in college or were supposed to enter college during this period were affected. Since universities were closed and the Iran-Iraq war was escalating, the affected cohorts were drafted into war. Using a regression discontinuity design, this paper documents the causal effect of this temporary elimination of supply of higher education on college attainment rates of these cohorts. The results show that Cultural Revolution reduced men's college attainment rate by only about 10 to 12 percent (1.4 to1.7 percentage points) and had no effect on women's. This corroborates the second argument discussed above as it shows that demand for education is crucial in educational attainment. This conclusion has important implications for higher education policy in both developing and developed countries.

In particular, it suggests that raising demand for education should be an important part of education policy.

This study also uses this exogenous decline in college attainment rate, caused by the Cultural Revolution, to estimate the return to college education for men. Although OLS estimates show that individuals with college and above education have 83% more hourly wages than others, the IV estimates precisely show a 121% return on hourly wages as a result of earning college and above education. Souri (2004), in his Ph.D. dissertation supervised by Djavad Salehi-Isfahani at Virginia Tech, has used deviations from the trend in college attainment rate for cohorts affected by the Cultural Revolution as an instrument for college attainment to estimate returns. This study, however, uses a regression discontinuity design in the first stage of the IV estimates. It argues that the first stage satisfies the exclusion restriction better than a deviation from the trend instrument. The results are also different.

The rest of this paper is organized as follows: Section 2 explains the Cultural Revolution in Iran. Section 3 is about the data. Sections 4.1, 4.2, and 4.3 present the correlational and causal evidence. Section 4.4 discusses the results and Section 4.5 reports the estimates on returns to education. The paper concludes in Section 5.

2. Cultural Revolution in Iran

The Islamic Revolution of 1979 in Iran has been considered a turning point in the recent history of the country and by some accounts, the Middle East. From the onset, its leaders considered this political revolution as a cultural and ideological revolution that aimed to establish a new paradigm in Iran to replace the notions of the Western culture (Mehran 1997). Reforming cultural institutions and particularly, the educational system was of high priority early on. Universities were of particular

importance as they were one of the main centers of political activity during and after the revolution. In the few months after the revolution, the transitional government protected most political freedoms. But eventually, the political elite thought of the political activities in universities as a major source of instability. In addition, soon after the revolution, the country was in an upheaval: a civil war and international crises. In the midst of these crises, the new political system which craved for domestic stability embarked on a consequential plan to close all institutions of tertiary education and halt any higher educational program in progress.[2]

Almost a year since the revolution, after a series of events from March through May of 1980, the Revolutionary Council ordered the closure of all universities on June 5, 1980.[3] This was the beginning of one of the most important educational projects in recent history of the country: the overhaul of the higher educational system that was called Cultural Revolution. To manage the process, on June 13, the Cultural Revolution Committee was established by the government. Its goals were to redesign the programs and curricula in universities in order to make them aligned with the ideologies of the revolution, and to select and train the right faculty to teach the new curricula. Particular attention was on programs in humanities and social sciences as their curricula were considered "un-Islamic."[4]

Closure of universities, that happened earlier on June 5, 1980, benefited the committee to implement its goals. Finally, on December 18, 1982, after thirty months of closure, the universities were re-opened. Before that, however, in the spring of 1982, those who needed less than 25 credits to graduate, were allowed to enroll. The first cohort of new students enrolled in fall of 1983.

---

[2] For a detailed account of the early history of the revolution see Bakhash (1984) and Hiro (1985).
[3] See Bakhash (1984) for a detailed account of events.
[4] This is still an important discourse in the political and social construct of the Islamic Republic.

The Cultural Revolution, however, did not end with the closure of universities. The Cultural Revolution Committee whose name was changed to the High Council of Cultural Revolution still exists and continuously updates textbooks and curricula among other activities. For simplicity, from now on, when we use the term Cultural Revolution, we refer to the period between 1980 and 1982 when institutions of higher education were closed.

Three cohorts of high school students were affected by the closure of universities: those who could enroll in universities in 1980, 1981, or 1982 but faced the closed door. No question in the data specifically asks whether the individual belonged to this group or not. In other words, no question asks whether the individual wanted to enter college in 1980, 81, or 82. But, with some approximation, one can argue that at least for men, this group consisted of individuals who were 18 in 1980, 1981, or 1982. This is because male high school graduates could try to enter college before age 19. After reaching age 19, they have to go to military service. This was particularly well-enforced at the time,[5] because the country was at war with Iraq and potential soldiers were in high demand.[6] Of course, after completing their service which was for two years, they could again try to enter college if they wanted to. But, they certainly missed the boat when they were 18.

Women, on the other hand, did not have this constraint and could have attempted to go to college even after they reached 19. But, women who were 18 in 1980, 81, or 82, and wanted to go to college but faced a closed door, had two options left: waiting until the schools re-open and apply to college, or marry off. These two options were relatively mutually exclusive, since after marriage and having children, the chance of going to college was potentially negligible. Considering the fact that the

---

[5] It is still very well-enforced.
[6] Abrahamanian (2008) compares this war , which was the longest war of the twentieth century, to World War I, in terms of military tactics, such as trench warfare, human wave attacks, and the use of chemical weapons such as mustard gas by the Iraqi government. The human wave attacks, in particular, required large body of soldiers and maximized demand for recruits.

median age at first marriage for women was about 17 at the time[7], women who were affected by the Cultural Revolution were under significant pressure to marry off, especially that it was not clear when the schools were supposed to re-open. In this paper, we estimate the effect of university closure on educational attainment of both male and female cohorts who became 18 in 1980-82.

The cohorts who were 18 in 1977, 78, and 79 were in college when the Cultural Revolution happened. So their education remained incomplete. They needed to wait until after the re-opening in order to complete their education. Therefore, they were potentially less likely to finish their education or attain higher levels of education (e.g. graduate school). We cannot identify partial vs. full completion of college in the data. Partial and full completion are coded similarly. Therefore, it is not possible to estimate the effect of Cultural Revolution on college completion for these cohorts.

3. Data

Two groups of datasets are used in this study. The first group contains the 2% censuses of 2011 and 2006 combined to estimate the impact of Cultural Revolution on educational attainment. But, because the censuses do not contain hourly wages, Household Expenditure and Income Surveys (HEIS) from 2006 through 2015 are used to estimate the return to education (using Cultural Revolution as an instrumental variable for education.) All these datasets are collected by the Statistical Center of Iran (SCI). Both censuses are available on IPUMS International website[8] and the HEIS datasets are on SCI website.

The 2% nationally representative samples of censuses have close to 2.84 million individuals and about 765,000 households combined. At the individual level, the data contains demographic

[7] Median age at first marriage was 17 years for all women aged 10 to 49, according to the Demographic and Health Survey of 2000.
[8] IPUMS website address is https://www.ipums.org/ and the SCI website address is http://amar.org.ir/english

information, such as gender, age, relation to the head of the household, birth year and month, education (in categories), employment and job characteristics, marital status, number of children ever born, migratory status, reasons for migration, and more. The largest sample that is used in this study consists of individuals born between 1931 and 1984. Summary statistics of all variables for this sample are reported in panel A of Table 1.

HEIS have been collected annually since 1984 and contain disaggregated expenditure and income as well as some household demographic data. But it has been only since 2006 that they contain hours worked. In fact, they are one of the few datasets in Iran that have hours worked. This lets us calculate wages and estimate the return to education. Each survey has about 15,000 to 30,000 households. Rural areas are over sampled. The surveys are generally used to estimate macro-level variables such as inflation and unemployment. After Censuses, they are the most important and regular project of Statistical Center of Iran. HEIS, however, have limited demographic data. For example, although the Censuses have year and month of birth, HEIS have only age. Therefore, birth years are constructed based on age. The summary statistics of the variables used are reported in panel B of Table 1.

An important issue with all Iranian data (censuses or HEIS) is that education is identifiable in levels: primary, middle school, high school, college, and graduate school. Years of education, grade, and whether the level was fully or partially completed are not correctly identifiable. Therefore, our outcome variable can best be defined as a dummy variable that shows whether the individual has partially or fully completed college or graduate school, or not. We call this dummy variable "College and above" in the rest of this paper.

One technical point about the data is that dates are recorded according to the Persian Calendar which is a solar calendar but different from the Gregorian one. The Persian new year, which has

been used in Persionate countries (Iran, Afghanistan, and Tajikistan) for centuries, starts on the spring equinox (the first day of spring) which is usually March 20th and ends a day before the spring equinox in the next Gregorian year. For example, the year 1390 in the Persian Calendar starts on March 20th, 2011 and lasts until March 19th, 2012. Therefore, it approximately overlaps the last nine months of 2011 and the first three months of 2012.[9] For simplicity, in presenting the results, we correspond each Iranian year to the Gregorian year with which it has the most overlap (i.e. nine months). For instance, 1390 corresponds to 2011. This is only to make the presentation of the results simpler and clearer and is not used to perform the econometric analysis. The reader should just note that wherever a Gregorian year is mentioned in the text or figures, the corresponding Persian year is the true reference in the data.

4. Methodology and Results

As discussed in Section 2, cohorts who were 18 in 1980, 81, or 82 wanted to enter college but faced a closed door, because of Cultural Revolution. They were directly affected by it. At the same time, cohorts who were in school when the Cultural Revolution started in 1980, were partially affected as their education remained incomplete. Those include cohorts who were 18 in 1977, 78, and 79. Based on these, birth cohorts of 1962, 63, and 64 were fully affected and birth cohorts of 1959, 60, and 61 were partially affected.

---

[9] Each month in the Persian Calendar matches the zodiac signs. This is because it is devised by mathematicians and astrologers of the 11th century upon request from the king, Jalal-e-Din Malek Shah. In his name, it is also called the Jalali calendar. Although widely in use since that time, with the advent of modern era in Iran, on February 21, 1911, the Iranian parliament adopted it as the official calendar. On March 31, 1925, the parliament made small modifications in the number of days in each month to make the calendar simpler. Based on that, the first six months have 31 days and the next five months have 30 days. The last month has 29 days, or 30 days in leap years. The New Year time corresponds to the spring equinox (the time the distance between the earth and the sun is averaged) which happens at a specific time on March 20th or a close date. Therefore, the specific time an Iranian New Year starts can be calculated. For instance, the Year 1395 starts at 8:04:30 pm on March 20, 2016 (Tehran time). The months in the calendar are named after the ancient Persian names for the zodiac signs. The calendar is also used in Afghanistan but the zodiac names are in Arabic, the scientific language of the period when the calendar was devised.

In this section, we try to identify the impact of Cultural Revolution on education of these cohorts. First, we demonstrate the descriptive evidence on deviations from the trend. Later, we employ a regression discontinuity design to estimate the causal effect by taking the difference between cohorts who were affected with slightly younger cohorts who were not.

4.1 Deviations from the Trend

The first step is to examine the college attainment rates across cohorts. Figures 1(a) and (b) describe these rates for men and women, respectively. The horizontal axis has the birth cohorts starting from 1931 through 1981 and the vertical axis shows the share of the population who had college and above education, i.e. undergraduate or graduate levels. Each dot represents the share of those who had college and above education in a birth cohort. As explained in the data section, this includes partial completion of each level as well. Since the data set is collected in year 2011 and 2006, these cohorts were between 22 to 80 years old at the time of data collection. Therefore, they should have been able to at least enter college by their age at the time of the data collection. The red curve represents the local polynomial fit and the gray curves are the 95% confidence intervals. The cohorts that are fully or partially affected by the Cultural Revolution are marked with dark and light gray bars, respectively.

As depicted in Figures 1(a) and 1(b), there is an upward trend in college attainment rates as one moves from older to younger cohorts. This has been documented in the past in Salehi-Isfahani (2005) and Majbouri (2010) among others. Eyeballing Figure 1(a), one can deduce that the share of men with college or more education in cohorts who were fully affected by the Cultural Revolution (i.e. 1962, 63, and 64 cohorts) deviates downwards from this trend. The rate is approximately between 1 to 2 percentage points lower than if the pre-Cultural Revolution trend was enduring for these male cohorts (14% vs. 15.5%).

9

Interestingly, looking at Figure 1(b), it is not clear (visually) whether there was any decline in the share of women with college and above education among the affected cohorts. It seems that the trend continues upward (without any discontinuity) even for cohorts who were part of the Cultural Revolution. The rise in female college attainment is faster than males. This is the trend that made female students make up two thirds of college enrollments in the 2000s.

A more careful consideration than a visual observation involves estimating the deviation from the trend for the affected cohorts and testing for its statistical significance. Simply, one assumes various functional forms for the trend and estimates the deviations from those functions at the time of the Cultural Revolution. Appendix B explains and reports these results which show that deviation from the trend for the male cohorts who graduated from high school and faced the closed university doors is about 1.2 percentage points. For women, however, this deviation is sensitive to the specification and switches signs (negative to positive). Note that although the deviation from the trend for men (the regressions in Appendix B) is unique to the 1961-64 cohorts and is consistent with what one expects of the effect of Cultural Revolution, it only shows the correlation between the Cultural Revolution and college attainment. It happens at the same time as the Cultural Revolution, but, obviously, does not show a causal effect.

4.2 Regression Discontinuity Using Birth Years

In order to document the causal effect of Cultural Revolution on attainment rates, a regression discontinuity design is employed. Universities were re-opened for cohorts who were born after 1964, but they were closed for the three cohorts who were born just before 1964. Because of Cultural Revolution one may expect a discontinuity in attainment rates at the 1964 cohort.

Figures 2(a)-(c) fit three different local polynomials (first order, second order, and local smooth polynomial respectively) before and after the 1964 birth cohort, for the sample of men. Figures 2(d)-

(f) report the same for the sample of women. The horizontal axes represent the birth year cohorts measured as the number of years since the threshold which is 1964. The vertical axes show the share of those who partially or fully completed undergraduate or graduate schools. There seems to be a sharp discontinuity at 1964. Cohorts who were born just before 1964 have lower attainment rates than those just above it. It does not however seem that the same results apply to women. The discontinuity for women is positive and small in a linear regression (Figure 2(d)), but negative and small in the others. They are not conclusive visually.

Figure 2 offers suggestive evidence of discontinuity for men and corroborates that a regression discontinuity (RD) is a natural solution to estimate the causal effect. The running variable in this RD design is birth year which is discrete. There are only a bit over 50 birth year cohorts that have enough observation and were old enough in 2006 or 2011 to attend college; those are birth cohorts between 1931 and 1984. The few discrete values of the running variable means that recently developed optimal bandwidth choice methods and local linear estimations cannot be employed and a standard classic regression should be used. But, for robustness check, various bandwidths can be examined in this standard regression. Table 3 shows the result of the following standard regression discontinuity design:

$$C_i = \alpha + \beta D_i + \sum_{l=1}^{L} \gamma_l (B_i - 1964)^l + \sum_{l=1}^{L} \delta_l D_i (B_i - 1964)^l + u_i, \quad L = 1, 2, 3 \quad (1)$$

in which C_i is a dummy equal to one if individual i has partially or fully attended undergraduate or graduate schools. B_i is the year individual i was born (i.e. birth cohort). $B_i - 1964$ is the running variable in the regression discontinuity design.[10] L is the degree of the polynomial control for the

---

[10] 1964 is the cutoff point based on the Gregorian Calendar. But the birth year is reported in Persian Calendar format in the data. 1964 is equivalent to 1343 in the data. Therefore, B_{is} is differenced with 43. Following Cattaneo et al. (2017), the birth year 43 is changed to 42.999 to make sure it falls on the other side of the threshold.

running variable. D_i is a dummy equal to one if individual i was born in or before 1964 and zero otherwise. The coefficient of D_i, i.e. β, shows the local average treatment effect of the Cultural Revolution. Lee and Lemieux (2010) suggest that when the running variable is discrete, standard errors should be clustered at the discrete values of the running variable. Following this, all regressions include cluster-robust heteroskedastic standard errors. But if we do not cluster standard errors or cluster them based on the 60 geographic regions, we get smaller standard errors and hence, more statistically significant estimates of the effects.

The left and right sections of the top panel in Table 3 report the local average treatment effect, β, for men and women, respectively. Fifteen birth cohorts before and after the threshold (i.e. 1964) were included in these regressions. In other words, the sample contains 1950 through 1979 birth cohorts. All specifications (first, second, and third degree polynomials), reported in Columns (1)-(3), show a negative and statistically significant effect of Cultural Revolution on attainment rates for men.

Similar to Figures 2(d)-(f), the results for women, i.e. Columns (4)-(6), depict little evidence of discontinuity. There is only a statistically significant discontinuity for the first order polynomial setting which vanishes and switches sign in the second and third order specifications. Interestingly, like Figure 2(d), the coefficient is positive in the first order setting implying that cultural revolution increased educational attainment; an odd result. Moreover, Figures 1(b) and 2(d)-(f) show that the trend in college educational attainment for women is not linear but quadratic. So the correct specification is not the first order polynomial. All these imply no effect of cultural revolution on women's college attainment rates.

Since the optimal bandwidth choice methods cannot be implemented on the few discrete values of the running variable, the birth year, one needs to check the robustness of the RD using various

bandwidth sizes and classical regression. The bottom panel shows the the local average treatment effect (for the three polynomial orders) when we change the bandwidth of the sample from 15 years to 8, 10, 12, and 20 years before and after 1964. Each row represent a bandwidth size and each column is a polynomial order (first, second, and third). For men (the left section of the bottom panel), the first and second order polynomial results are always statistically significant (mostly at 1% level). The third order polynomial is also statistically significant when the bandwidth is 12 years (or 15 years as reported in the top left panel of the table). This could be because of overfitting for higher than second order polynomials, because the trend in Figure 1(a) becomes quadratic as we get closer to the cut-off point.

As the right section of the bottom panel shows, for women, there is no discontinuity in almost all specifications and bandwidth choices. In the two out of twelve instances that the coefficient is significant (first and third order polynomials for ±20 years sample), it takes opposite signs (positive for first order and negative for the third). There is no evidence that the Cultural Revolution affected women. This could be because women who were likely to go to college at the time – a small fraction of women – were quite different from others and had such a high demand for college education that they would have gone to college anyway. They were such a selected group of women that went to college as soon as opportunity arose later.

4.3 Regression Discontinuity Using Birth Months

In Iran, like many other countries, children who have reached age six can enroll in primary schools. More specifically, those who reach age six before the new educational year begins (September 23rd of every year), can start their school in the same year. But those who reach age six after the educational year starts, should wait and enroll in the following educational year. In other words, if someone is

born in July 1964, she becomes six years old in July 1970 and can start her first grade in September of the same year, i.e. 1970 (when the 1970-71 educational year begins). But someone who is born in December of 1964 reaches six in December of 1970 and therefore, can go to school in September of 1971. She is still six then, but in three months she will be seven. So there is a discontinuity in the year that an individual starts school based on her month of birth.

In Iran, the educational year starts on the first day of *Mehr* (mostly, September 23rd)[11], the seventh month of the year. In practice, those who were born in Mehr and in some cases those who were born in *Ābān*[12] (the eighth month) still get to go to school in the same year as those who were born earlier. This is because many parents feel it is unfair that their children have to wait one more year in order to go to school only because of being born a few days late. Therefore, they do all they can to enroll their children in the same year as those who were born before Mehr and they are successful in most cases. Hence, in practice, Ābān is a better cut-off point for the birth month of children who start school in a year.

Now, consider the fact that every Iranian male should go to military service in the month that he becomes 18 years old, unless he is in school. A man who was born a month before Ābān of 1964 could attend first grade in 1970. But a second man who was born a month after Ābān of 1964 could only go to first grade in 1971, i.e. a year later than those who were born before Ābān of 1964. So, eleven years later, in December of 1982, when the Cultural Revolution was about to end and the universities were about to re-open in the following year, the second man was still in his last grade of high school while the first man had had finished high school a few months before and had already been drafted. Hence, the second man had the opportunity to enter college the next year, unlike the

---

[11] Mehr is pronounced /'mehr/ (using Meriam-Webster pronunciation keys) and corresponds to Libra, the seventh zodiac sign. As mentioned in footnote #7, each month in the Persian Calendar corresponds to an astrological sign.
[12] Ābān is pronounced as /'äb'än/ (using Meriam-Webster pronunciation keys) and corresponds to Scorpio, the eighth zodiac sign. See footnote #7 for more information.

first man. The first man was also unlikely to try to enter college during the time that he was serving in the war and maybe even afterwards. As a result of these, there should be a discontinuity in college educational attainment at the birth month of Ābān of 1964 that was caused by the Cultural Revolution.

Figures 3(a)-(c) and 3(d)-(f) illustrate this discontinuity for men and women, respectively. The running variable on the horizontal axes is the month of birth measured as the number of months since the threshold which is Ābān of 1964. Therefore, months before the threshold take negative values and those after it take positive ones. The vertical axes show share of college educated. Men born just before the threshold seem to have less likelihood of having tertiary educational attainment. Based on the results in Table 2, one expects to see no discontinuity at the threshold for women. Figures 3(d)-(f) support that.[13]

The running variable, month of birth since Ābān of 1964, is a discrete variable. But because it can take over 600 values, we can implement the latest developments in RD estimation, find the optimal bandwidth, and estimate a local polynomial around the threshold (for more information, see Calonico et al., 2017; Cattaneo et al., 2014a,b; and Cattaneo et al., 2017). Table 3 reports the results for such estimations. There are two major methods of optimal bandwidth choice in the literature: 1) Mean Square Error (MSE) and 2) Coverage Error Rate (CER). Each of these methods can find the same optimal bandwidths above and below the threshold (we call it R.D. from now on) or different optimal bandwidths above and below the threshold (we call it Two-sided from now on). So four different optimal bandwidths choice methods can be implemented: MSE-R.D., MSE-Two Sided, CER-R.D., and CER-Two Sided. We estimate results for all these choices in Table 3, for robustness check. The data extends from 1931 to 1984 birth cohorts. 1984 is chosen because we want the

_____

[13] Figures 3(a)-(f) are drawn with the `rdplot` command in Stata.

youngest age in the sample to be about 22 (the 1984 cohort is about 22 years old in the 2006 census). 1931 is chosen because we want to have substantial number of observations in the oldest cohort (the 1931 cohort is about 80 years old in the 2011 census). The threshold is Ābān of 1964. Therefore, there are about 20 birth cohorts above the threshold and 33 below it. Since the data is not symmetric around the threshold (20 years above and 33 below), it is better to find separate optimal bandwidths for above and below the threshold. Therefore results for two-sided approaches (MSE–Two Sided and CER–Two Sided) are more relevant.

The top panel of Table 3 reports the discontinuity in educational attainment conditional on a first, second, and third order local polynomial of the running variable when the optimal bandwidths choice method is MSE-Two Sided. The results support those found in Table 2 using a standard classic regression discontinuity and the birth year as the running variable. For men, there is a discontinuity ranging from 1.2 to 1.7 percentage points in college attainment rates that is robust even if one uses third order local polynomial estimation. The results for women are all statistically insignificant and close to zero (0.1 to 0.2 percentage point).

The bottom panel reports the discontinuity using other optimal choice methods: CER-Two Sided, MSE-R.D., and CER-R.D. Each row reports one of these methods and each column presents an order of the polynomial (first, second, and third). Interestingly, all results for men are similar and statistically significant regardless of the order of the local polynomial used or the optimal choice method implemented. The discontinuity ranges only from 1.1 to 1.9 percentage points and most are between 1.5 to 1.7. All coefficients for women are similar to those found in the top panel as well: statistically insignificant and small.

4.4 Discussion

During Cultural Revolution, all universities were closed for 30 months. Three cohorts of young men faced immediate draft into the longest war of the twentieth century. Yet, they are only slightly less college-educated than the following cohorts. This is indeed puzzling.

One explanation is that the affected cohorts disproportionately went to war and died. Fewer of them survived and that is the reason for their relatively high college attainment rate. In other words, they needed to have fewer people attending college to achieve the same attainment rates. If this is true, we should see a discontinuity in the size of cohorts at the threshold. Figures A2 through A4 in the appendix plot the number of observations in each birth year and birth month cohort and fit multiple polynomial functional forms to the data. The figures present little evidence of discontinuity in the size of cohorts at the threshold. If any, the affected cohorts by the cultural revolution (cohorts just below the threshold) have more observations than those above it.

Table A1 reports the results of estimating Equation (1) when the dependent variable is the number of observations in the birth year cohorts. This table is in the same format as Table 2. The results for the bandwidth of ±15 years are reported in the top panel and the bottom panel depicts statistical significance levels of the discontinuity in cohort size for various bandwidths. The results for men show that there is either no statistically significant discontinuity in the cohort size or when there is, the affected cohorts by the cultural revolution are larger. Overall, there is no evidence of a smaller cohort size for the affected cohorts. The discontinuity for cohort size is either statistically insignificant or positive whenever significant. Therefore, the reason behind the relatively higher than expected attendance rates for the affected cohort was not that they survived less. If anything, they actually survived more. For women, the discontinuity is almost always statistically insignificant and switches signs depending on the specification.

One can use the birth month cohorts and implement a robust RD design using various optimal bandwidth choice methods. Table A2, which is in the same format as Table 3, reports the results. For both men and women, there is no discontinuity in the cohort size regardless of the optimal bandwidth choice and the polynomial order of the running variable. These results are all rejecting the above hypothesis that affected cohorts achieved a relatively higher than expected college attainment rate (despite the Cultural Revolution), because they were smaller than non-affected cohorts (and could have fewer college educated individuals to achieve such rates.) According to the best estimates, the total number of casualties during this war (on the Iranian side) ranges from about 200,000 to 600,000 (Hiro, 1991; Abrahamian, 2008; Bercovitch and Jackson, 1997; Clodfelter, 1991; Palmowski, 1997; Mikaberidze, 2011), and these casualties were spread across many birth cohorts involved in the war. Therefore, the casualties from the war should not have much of an impact on the size of the affected cohorts relative to others.

Another explanation for the small discontinuity in the educational attainment rate is that the affected cohorts went to war. These veterans of war received special quotas to enter university after their return. These quotas gave them a higher chance to enter college than the regular college applicants. But, even with these quotas in place, the small effect of the Cultural Revolution is a testament to high demand for college education in Iran (at least for the affected cohort.) The demand is large enough that even after such tumultuous years and thirty months of university closure, these cohorts thrive to go to school.

4.5 Returns to College Education

One may want to use this discontinuity in education levels as an instrument for college attainment and estimate the market return to college education. Unfortunately, wage rates (or income levels) are not reported in the census data. Instead, HEIS data, which contain hours worked since 2006, can be

used to estimate hourly wages for salaried workers. Combining 2006 through 2015 HEIS data, we obtain a sample of over 100,000 salaried workers. The summary statistics for this sample is reported in panel B of Table 1.

We estimate a simple Mincer equation for log of hourly wages in which there is only one education variable: College & above. It is a dummy variable equal to one if the individual has partially or fully completed college or graduate school and zero otherwise. College & above is instrumented with Cultural Revolution. In other words, the regression discontinuity, i.e. Equation (1), is the first stage for this two stage Mincer estimation. All the exogenous variables in this regression discontinuity are based on date of birth. In the HEIS data, however, the year or month of birth are not recorded. Therefore, we create the birth year using the age variable. Birth month, however, cannot be generated. Hence, we only use birth year as the running variable. Similar as in the top panel of Table 2, we use the sample of birth cohorts ±15 years around the threshold (i.e. 1950 through 1979).

Table 4 reports the result of estimating the Mincer equation by OLS and IV method in which Equation (1) is the first stage. The first column reports the OLS regression and the next three columns have the IV estimations. One can use three orders of polynomials of the running variable in the first stage. The results for each are reported in Columns (2) through (4). All standard errors are corrected for heteroscedasticity and correlation within the discrete values of the running variable (i.e. birth year).

The OLS estimate shows that those who have college and above education have 83% higher wages than those without it. The IV estimates, however, are about 121%. Interestingly, changing the order of the polynomial has little impact on the size of the returns (they range from 121% to 122%). That is people with college and above education, on average, earn over twice as much as people who have less education. The fact that IV estimates are larger than OLS estimate means that measurement

error is larger than omitted variable bias in the estimation of return to education. Note that these estimates are local average treatment effect. In other words, these are the estimates for individuals in birth cohorts around the threshold. Nevertheless, there is little reason to assume the return is substantially different for other birth cohorts.

The Kleibergen-Paap rk Wald F statistics are substantially larger than the Stock-Yogo critical values for weak instruments. The critical values are about 16 to 19 while the test statistics are in several hundreds. This means that the regression discontinuity provides a strong first stage for the education variable.

5. Conclusion

The contemporary history of Iran is fascinating and puzzling. Fascinating as over a few months the monarchy collapsed and new institutions dominated the political power. The ruling class, overnight, were replaced with an elite class that has never had such political power in the past millennia. In the first two years, Iran was in upheaval. It experienced a civil war and the large migration of the educated to diaspora. In two years, Iran faced the biggest threat to its existence: the invasion from Iraq. This invasion turned into the longest war of the twentieth century.

In the midst of these crises, the Iranian government decided to close all of its universities for thirty months. The puzzling phenomenon reported in this paper is that cohorts who were affected by all these tumultuous upheavals, particularly the closure of universities were almost as likely to have been college educated. This is a testament to high demand of education in Iran that in fact shows itself in the rapid rise in average years of education during and after the Iran-Iraq war. Human capital investment has been a corner stone of the Iranian household after World War II and it has remained one of the resilient features of the economics of household in Iran.

Further anthropological and sociological research is necessary to understand how these affected cohorts gained the education they wanted years after they left high school.

References

Abrahamian, Ervand (2008). *A History of Modern Iran* (3rd print ed.). Cambridge, U.K.: Cambridge University Press.

Bakhash, Shaul (1984). *The Reign of the Ayatollahs: Iran and the Islamic Revolution*. New York: Basic Books, p. 122.

Bercovitch, J., and Jackson, R. (1997). *International Conflict : A Chronological Encyclopedia of Conflicts and Their Management 1945–1995*. Congressional Quarterly Press, p. 195.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). "`rdrobust`: Software for regression discontinuity designs," *Stata Journal*, forthcoming.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). "Robust data-driven Inference in the Regression-Discontinuity Design," *Stata Journal*, 14, 909-946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295-2326.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2017). "A practical introduction to regression discontinuity," Monograph prepared for *Cambridge Elements: Quantitative and Computational Methods for Social Science*. Cambridge, UK: Cambridge University Press.

Clodfelter, Michael (2002). *Warfare and Armed Conflict: A Statistical Reference to Casualty and Other Figures*. Jefferson, NC: McFarland & Company, p. 1618–1991.

Hiro, Dilip (1985). *Iran under the Ayatollahs*. London: Routledge and K. Paul.

Imbens, G. and K. Kalyanaraman (2012). "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies,* 79(3): 933-959.

Lee, D. S. and T. Lemieux (2010). "Regression discontinuity designs in economics." *Journal of Economic Literature*, 48(2): 281-355.

Majbouri, Mahdi (2010). *Against the Wind: Labor Force Participation of Women in Iran*. Ph.D. thesis. University of Southern California.

Mehran, Golnar (1997). "Education in post-revolutionary Persia, 1979-95." In Yarshater, Ehsan (Ed). *Encyclopedia Iranica*, Vol. VIII, Fasc. 3, pp. 230-233. Also available online at http://www.iranicaonline.org/articles/education-xxiv-education-in-postrevolutionary-persia-1979-95 (accessed on Dec. 23, 2015)

Mikaberidze, Alexander (2011). *Conflict and Conquest in the Islamic World: A Historical Encyclopedia*. Santa Barbara, California: ABC-CLIO, p. 418.

Palmowski, Jan (1997). *A Dictionary of Twentieth Century World History*. Oxford, UK: Oxford University Press.

Salehi-Isfahani, Djavad (2005). "Human resources in Iran: potentials and challenges." *Iranian Studies* 38 (1), 117-147.

Souri, Davoud (2004). "Using the IV Method to Estimate Omitted Variable Models: An Application of the IV Method to Estimate the Rate of Return to Education in Iran," Ph.D. Dissertation, Virgina Tech.

Figures

Figure 1 – Proportion of men and women with college and above education
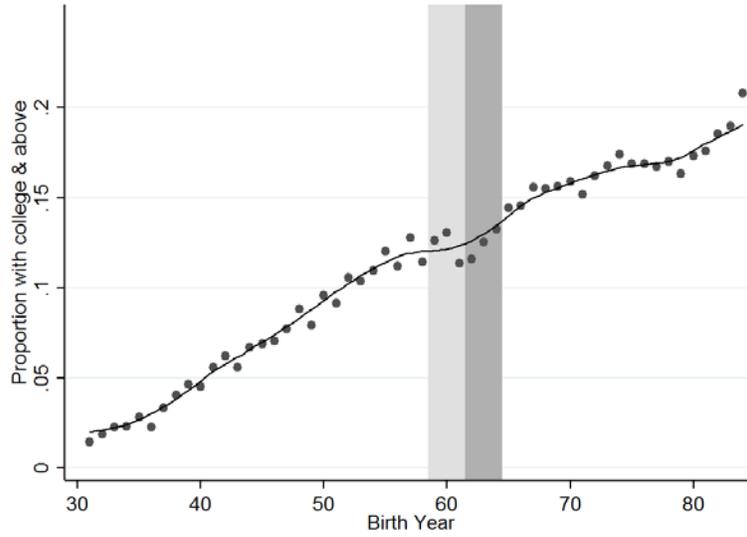

(a) Men
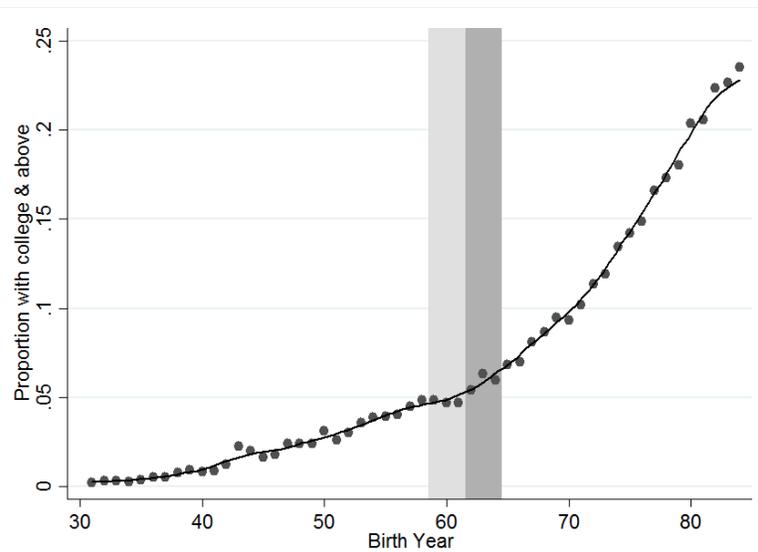
(b) Women

Note: College and above represents attainment of partial or complete undergraduate or graduate level education. Birth cohorts of 1962, 63, and 63 are shaded with dark gray as they reached age 18 when universities were close. Birth cohorts of 1959, 60, and 61 were in school just before the universities were closed. So their education could have remained incomplete. They are shaded with light gray.
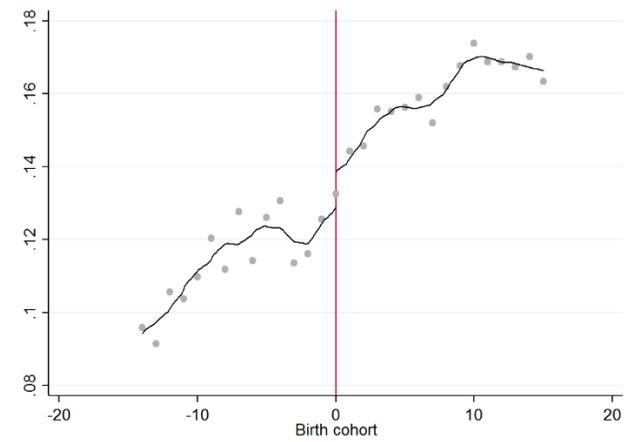
Figure 2 – Discontinuity in Proportion of men and women with college and above education across birth cohorts
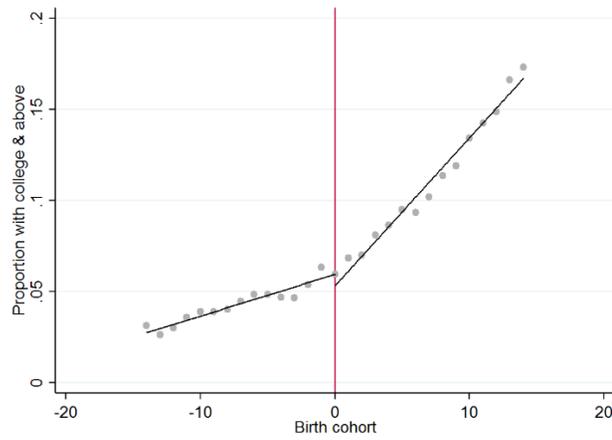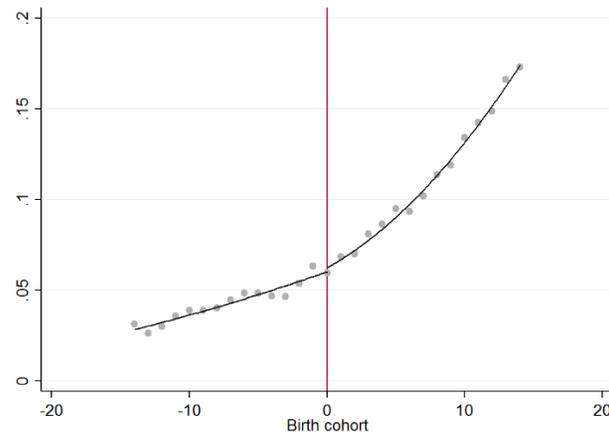(1950 through 1979 birth cohorts)
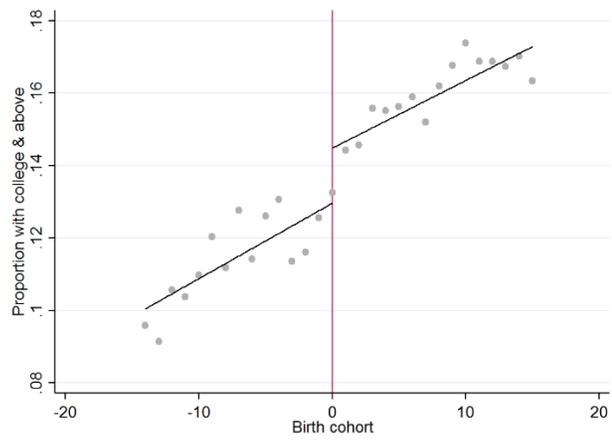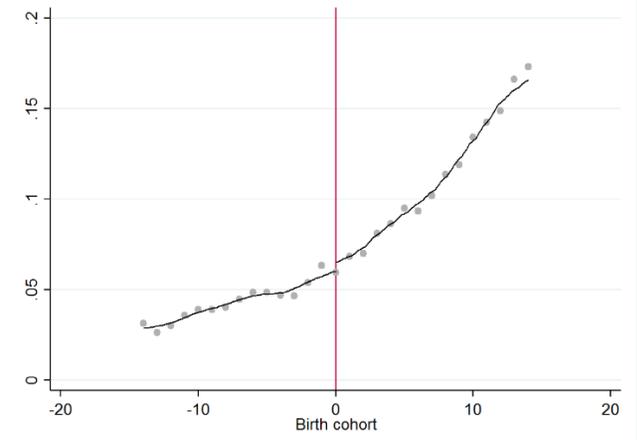

(a) linear - men
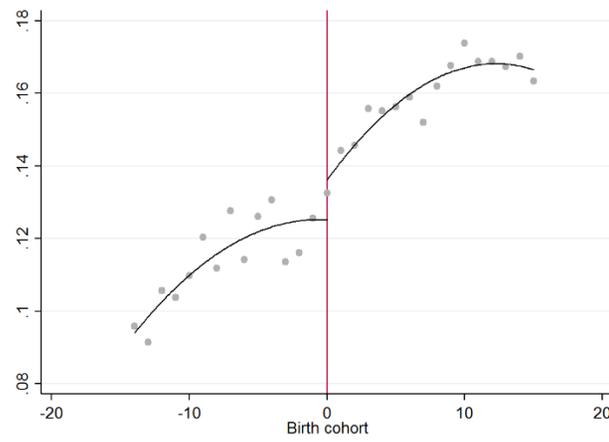
(b) quadratic - men

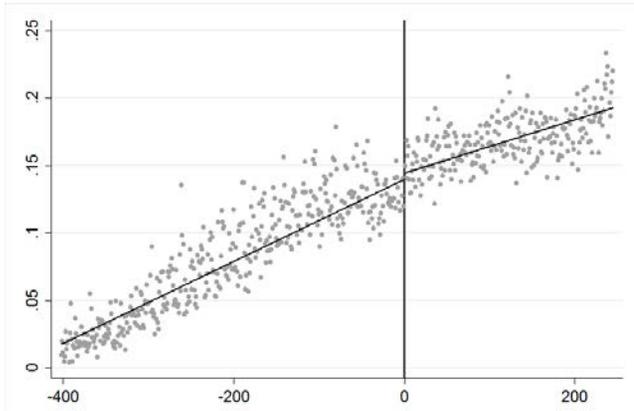(c) local polynomial smooth line - men

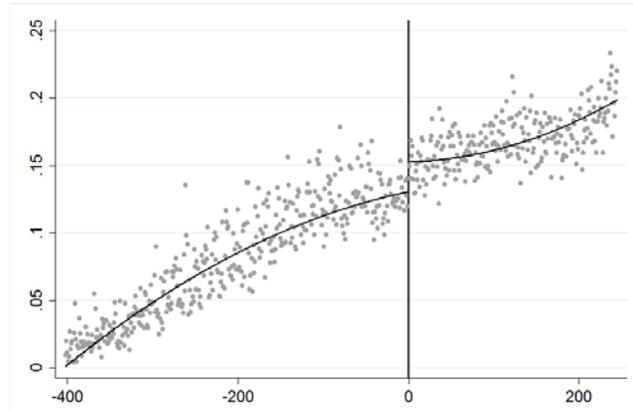(d) linear - women

(e) quadratic - women

(f) local polynomial smooth line - women

Note: The dots represent the proportion in each birth year cohort who partially or fully completed an undergraduate or graduate degree. The curves depict linear, quadratic and local polynomial predictions respectively.
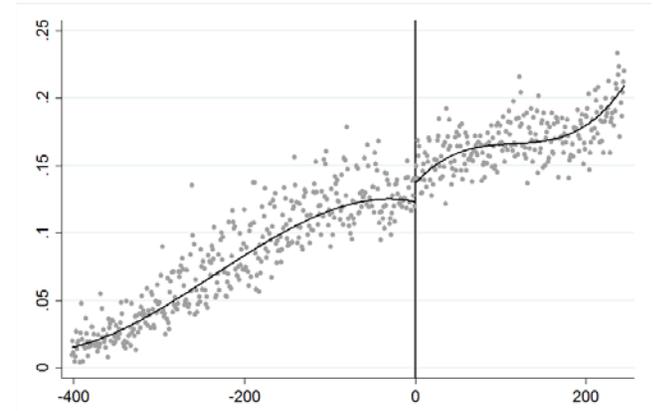
24

Figure 3 – Discontinuity in the proportion of men and women with college and above education (vertical axis) across monthly birth cohorts (horizontal axis)
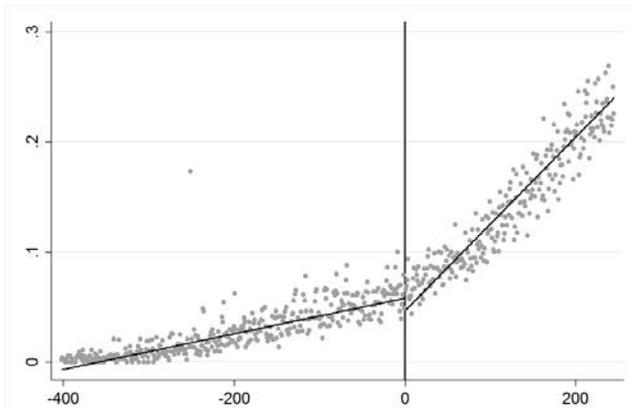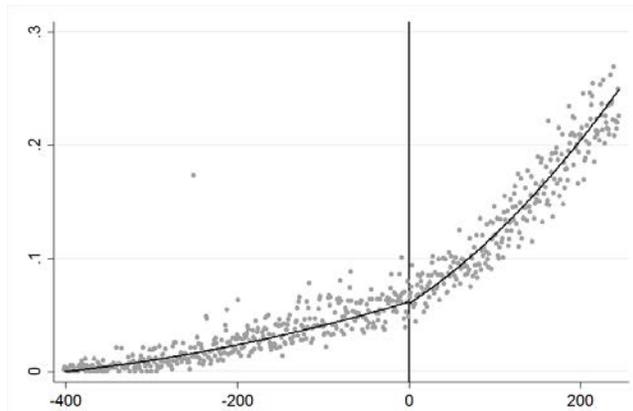(1931 through 1984 birth cohorts)

(a) local linear - men   (b) local quadratic - men   (c) local polynomial of the 3rd order - men

(d) local linear - women   (e) local quadratic - women   (f) local polynomial of the 3rd order - women
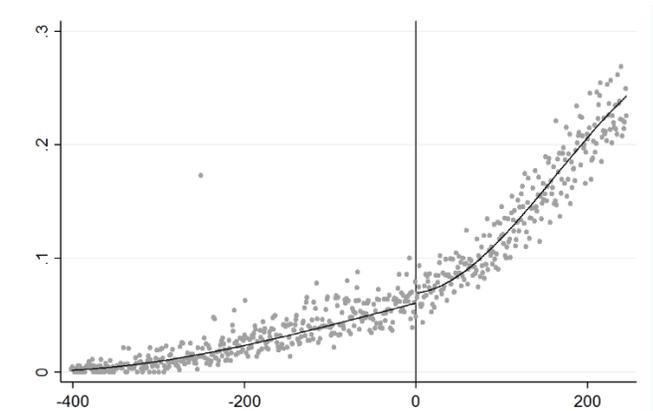
Note: The vertical axis represents the proportion of the birth cohort who partially or fully completed an undergraduate or graduate level education. The figures are drawn with `rdplot` command in Stata to find the optimal bin size. The dots represent the proportion in the bin who partially or fully completed an undergraduate or graduate degree. The curves represent first, second, and third order local polynomial predictions.

Tables

Table 1 – Summary Statistics

Variable	Observations	Mean	St. dev.	Min	Max
Panel A – 2006 and 2011 Censuses					
Female	1,470,257	0.50	0.50	0	1
Age	1,470,257	41.37	13.81	22	80
Literate	1,470,257	0.77	0.42	0	1
Primary school	1,470,257	0.24	0.42	0	1
Middle school	1,470,257	0.17	0.37	0	1
High school	1,470,257	0.22	0.42	0	1
College & above	1,470,257	0.13	0.33	0	1
Birth year	1,470,257	1967.09	13.61	1931	1984
Birth month	1,420,929	5.52	3.37	1	12
Panel B – Household Expenditure and Income Surveys, 2006-2015 (Male Wage Earners in 1950 to 1979 Birth Cohorts)					
ln(wage)	109,992	13.85	0.73	8.73	18.67
Age	109,992	41.90	7.87	27	66
Urban	109,992	0.55	0.50	0	1
College & above	109,992	0.12	0.33	0	1

Note: Female is a dummy equal to one if the individual is a female and zero otherwise. College & Above is a dummy variable equal to one if the individual has attended college or graduate school. Note that when Birth years are reported in Gregorian calendar but they are in Solar Hijri (Persian) calendar in the dataset. The sample for censuses includes individuals whose birth year is between 1931 and 1984. The sample for Household Expenditure and Income Surveys include only male wage earners in 1950-1979 cohorts. ln(wage) is the natural log of gross wages for wage earners. Gross annual salary is divided by the hours worked per week.

Table 2 – College Attendance and Cultural Revolution
Bandwidth: ±15 years, i.e. Birth cohort is between 1950 and 1979

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
D	-0.018***	-0.014**	-0.017***	0.009**	-0.003	-0.005
	(0.004)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)
Polynomial degree	1st	2nd	3rd	1st	2nd	3rd
Observations	485,418	485,418	485,418	482,351	482,351	482,351
Coefficient of D when the bandwidth is:						
±8 [i.e. 1957-1972]	-0.020***	-0.010**	0.003	-0.001	0.000	0.007
	(0.005)	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)
±10 [i.e. 1955-1974]	-0.016***	-0.018***	-0.000	0.002	-0.004	0.006
	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)	(0.006)
±12 [i.e. 1953-1976]	-0.016***	-0.016***	-0.013***	0.004	-0.003	-0.001
	(0.004)	(0.005)	(0.004)	(0.003)	(0.003)	(0.005)
±20 [i.e. 1946-1984]	-0.009*	-0.027***	-0.006	0.016***	0.001	-0.008*
	(0.005)	(0.007)	(0.007)	(0.004)	(0.003)	(0.004)

Note: Dependent variable is a dummy equal to one if the individual partially or fully completed college and above levels, and zero otherwise. Following suggestion by Lee and Lemieux (2010), robust-heteroskedastic standard errors are corrected for correlation within each birth year cohort, as this running variable (birth year) is a discrete variable. No other covariates except those in Equation (1) are controlled for. Results for each bandwidth selection in the bottom panel are reported in Table A2 in the appendix.
*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 3 – College Attendance and Cultural Revolution
Optimal Bandwidth Choice, Birth Month Cohorts

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
D†	-0.017***	-0.012*	-0.014**	-0.001	0.001	0.002
	(0.005)	(0.006)	(0.007)	(0.005)	(0.006)	(0.007)
Polynomial degree	1st	2nd	3rd	1st	2nd	3rd
Observations	710,131	710,131	710,131	710,798	710,798	710,798
Coefficient of D when the optimal bandwidth choice method is:						
Coverage Error Rate – Two Sided	-0.017***	-0.016**	-0.016**	-0.002	-0.001	-0.003
	(0.005)	(0.007)	(0.008)	(0.006)	(0.007)	(0.009)
Mean Square Error – R.D.	-0.012**	-0.011*	-0.012*	0.002	0.002	0.003
	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)	(0.008)
Coverage Error Rate – R.D.	-0.015**	-0.016**	-0.019**	0.0002	-0.002	-0.005
	(0.006)	(0.007)	(0.008)	(0.007)	(0.008)	(0.009)

Note: Dependent variable is a dummy equal to one if the individual went to college or above and zero otherwise. `rdrobust` command in Stata is used to estimate the optimal bandwidth and bin size. The top panel reports the coefficient of D when the optimal bandwidth method used is Mean Square Error – Two Sided (two different optimal bandwidths above and below the threshold). The bottom panel reports the same coefficient for other optimal bandwidth choice methods, for robustness check. Since the running variable is discrete, following Lee and Lemieux (2010), robust-heteroskedastic standard errors corrected for correlation within each birth month cohort (the running variable) are estimated and reported in parentheses.
 *** $p<0.01$, ** $p<0.05$, * $p<0.10$
† Optimal bandwidth choice method is Mean Square Error – Two Sided (two different optimal bandwidth choices above and below the threshold).

Table 4 – Returns to College Education for Men Using the Discontinuity
from the Cultural Revolution as an Instrument

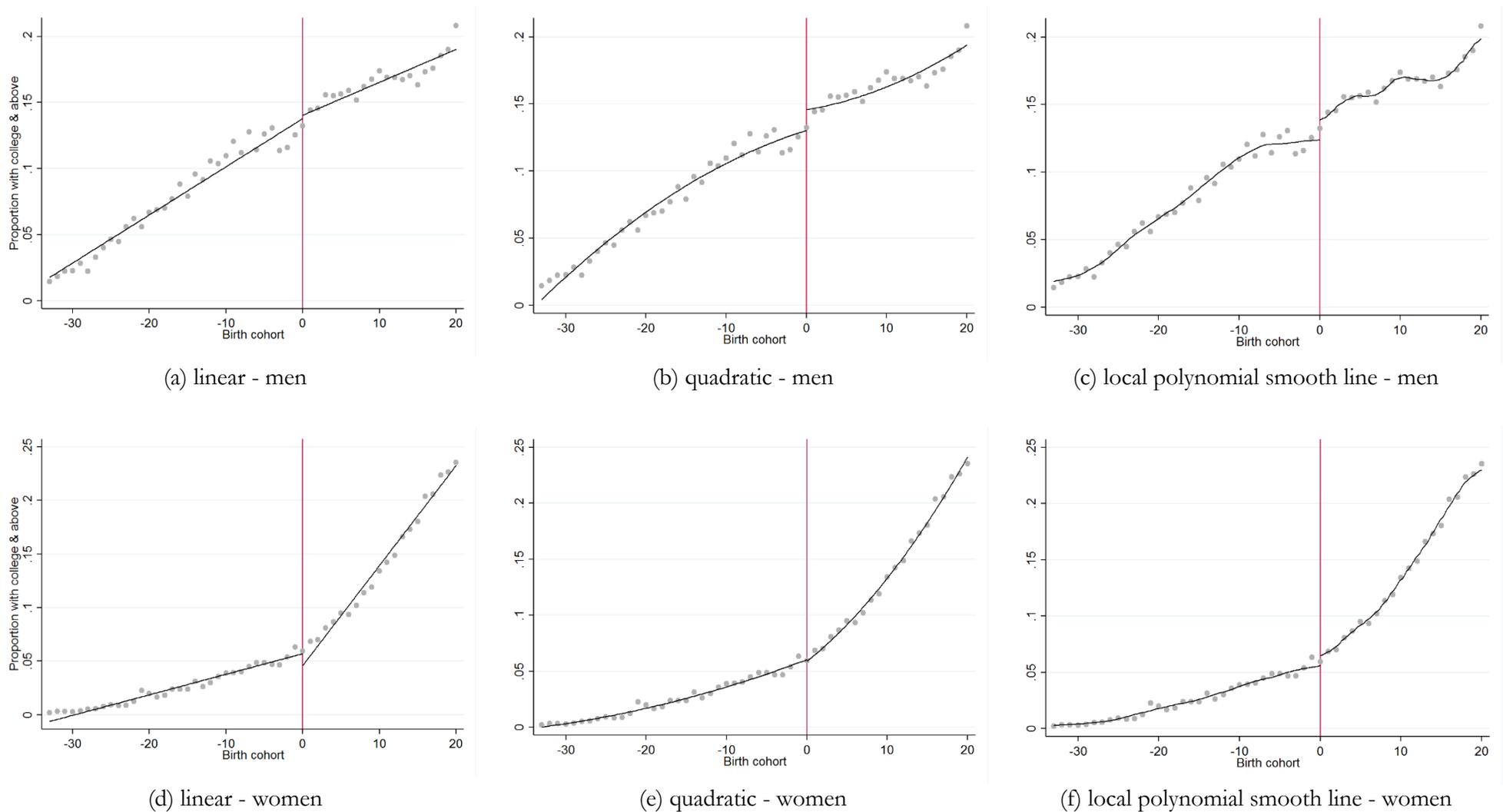|	All			
	OLS	IV		
	(1)	(2)	(3)	(4)
College & above	0.830***	1.222***	1.209***	1.211***
	(0.020)	(0.068)	(0.067)	(0.067)
Age	0.107***	0.101***	0.101***	0.101***
	(0.004)	(0.003)	(0.003)	(0.003)
Age$^2 \times 10^{-3}$	-1.170***	-1.084***	-1.087***	-1.086***
	(0.050)	(0.039)	(0.039)	(0.039)
Constant	11.393***	11.452***	11.450***	11.450***
	(0.093)	(0.080)	(0.080)	(0.080)
Polynomial order of the RD in the 1st Stage		1st	2nd	3rd
F-Statistic†		325.8	451.2	326.7
Observations	109,992	109,992	109,992	109,992

Note: Dependent variable is the natural log of average wage for every job category,
calculated from the Household Expenditure and Income Survey of 2010. The sample for all
these regressions consists of birth cohorts between 1950 and 1979 (i.e. ±15 years around
the threshold.)
*** p<0.01, ** p<0.05, * p<0.10
† Kleibergen-Paap rk Wald F statistic for weak instruments.

Figure A1 – Discontinuity in the proportion of men and women with college and above education across birth year cohorts
(1931 through 1984 birth cohorts)

(a) linear - men

(b) quadratic - men

(c) local polynomial smooth line - men

(d) linear - women

(e) quadratic - women

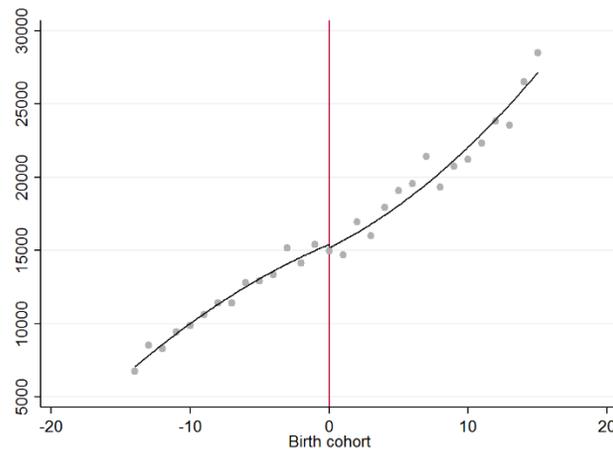(f) local polynomial smooth line - women

Note: The dots represent the proportion in each birth year cohort who partially or fully completed an undergraduate or graduate degree. The curves depict linear, quadratic and local polynomial predictions respectively.
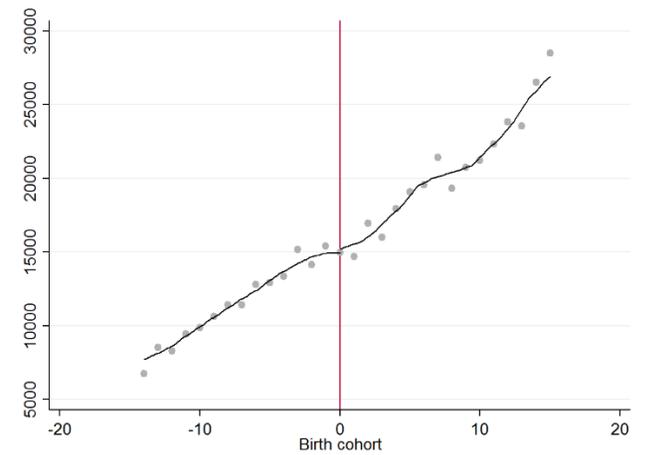
Figure A2 – Size of male cohorts and the Cultural Revolution – Birth year cohorts are between 1950 to 1979
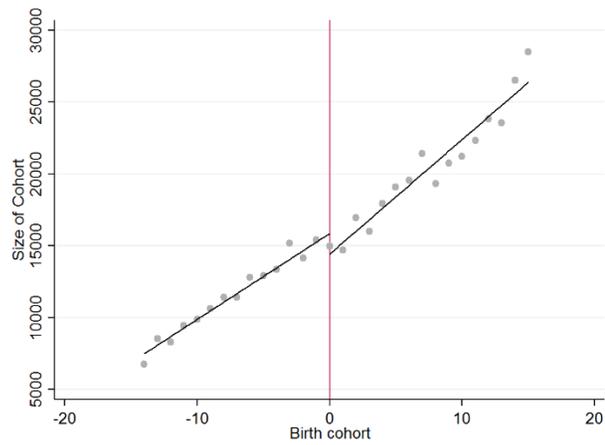
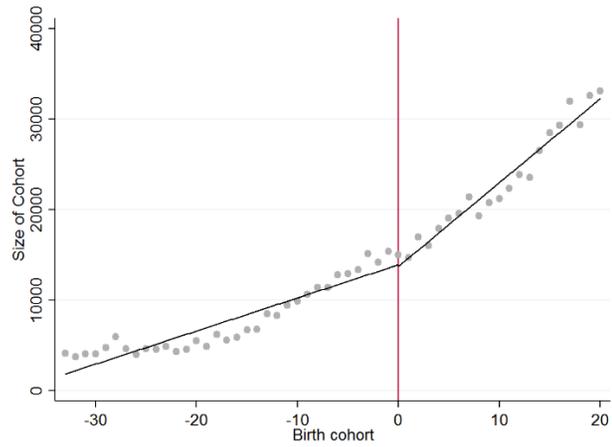(a) linear – birth year is 1950-79   (b) quadratic – birth year is 1950-79   (c) local polynomial – birth year is 1950-79

Note: The dots represent the proportion in each birth year cohort who partially or fully completed an undergraduate or graduate degree. The curves depict linear, quadratic and local polynomial predictions respectively.

Figure A3 – Size of male cohorts and Cultural Revolution – Birth year cohorts are between 1931 and 1984

(d) linear – birth year is 1931-84
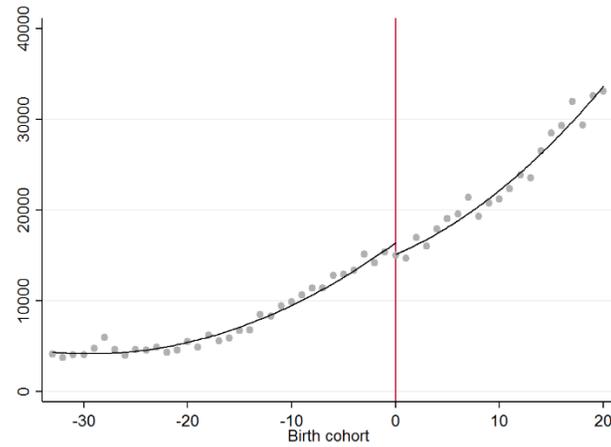
(e) quadratic – birth year is 1931-84
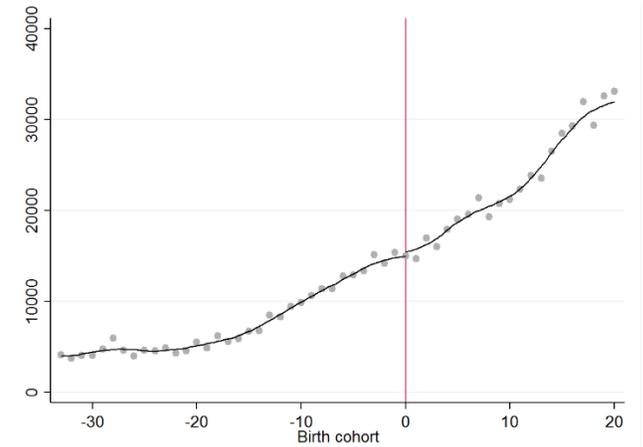
(f) local polynomial – birth year is 1931-84

Note: The dots represent the proportion in each birth year cohort who partially or fully completed an undergraduate or graduate degree. The curves depict linear, quadratic and local polynomial predictions respectively.

Figure A4 – Size of male cohorts (vertical axis) across birth month cohorts (horizontal axis)
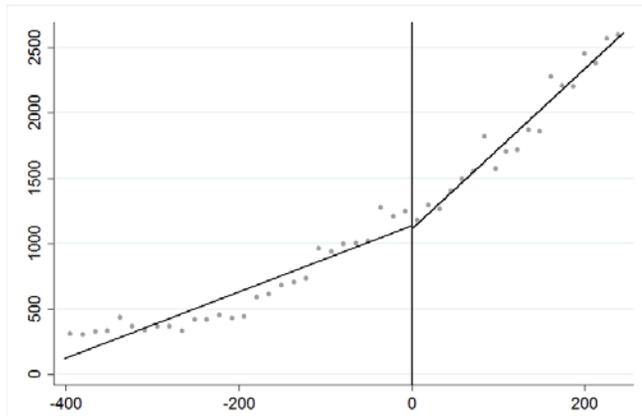
(a) linear

(b) quadratic

(c) polynomial fir of order 3

Note: The dots represent the proportion in each birth year cohort who partially or fully completed an undergraduate or graduate degree. The curves depict linear, quadratic and local polynomial predictions respectively.
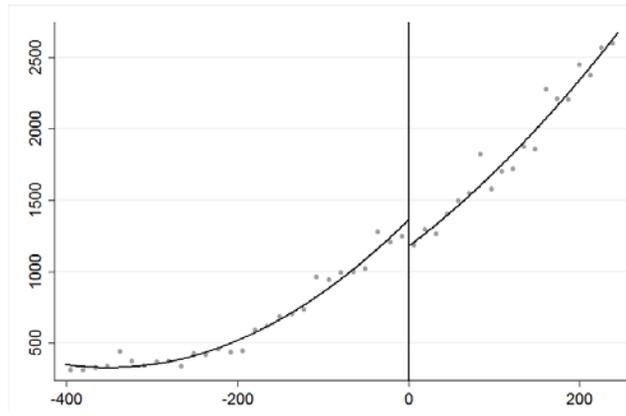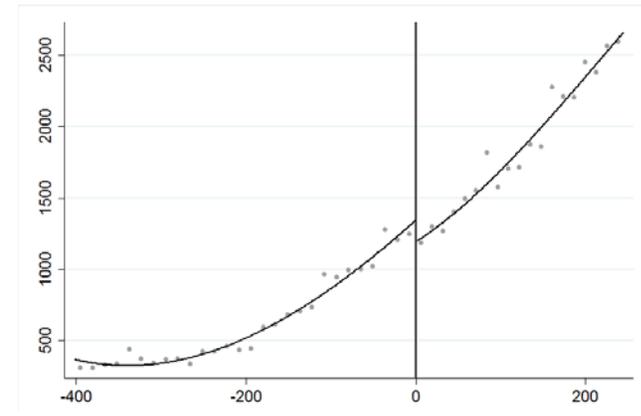
Table A1 – Cohort Size and Cultural Revolution
Bandwidth: ±15 years, i.e. Birth cohort is between 1950 and 1979

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
D	1,588**	54	2,546***	927	-1,008	1,341
	(623)	(1,016)	(792)	(682)	(1,168)	(801)
Polynomial degree	1st	2nd	3rd	1st	2nd	3rd
Observations	30	30	30	30	30	30
Coefficient of D when the bandwidth is:						
±8 [i.e. 1957-1972]	1,030	1,929**	-451	4	894	100
	(890)	(763)	(1,954)	(932)	(834)	(1,947)
±10 [i.e. 1955-1974]	694	1,788**	1,376	-193	597	961
	(700)	(828)	(1,145)	(781)	(927)	(1,113)
±12 [i.e. 1953-1976]	869	1,056	2,356***	134	-155	1,501**
	(611)	(953)	(613)	(676)	(1,088)	(663)
±20 [i.e. 1946-1984]	2,309***	575	195	2,245***	-767	-771
	(566)	(805)	(1,167)	(739)	(931)	(1,303)

Note: Dependent variable is the size of each birth year cohort. Robust standard errors are in parentheses. All statistically significant coefficients in the bottom panel (for various bandwidth sizes) are positive (for both men and women).
*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table A2 – Cohort Size and Cultural Revolution
Optimal Bandwidth Choice, Birth Month Cohorts (1931 through 1984)

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
D†	217.4	254.6	275.3	152.5	181.4	212.3
	(182.9)	(221.3)	(259.6)	(166.0)	(202.0)	(237.1)
Polynomial degree	1st	2nd	3rd	1st	2nd	3rd
Observations	648	648	648	648	648	648
Coefficient of D when the optimal bandwidth choice method is:						
Coverage Error Rate – two sided	249.1	288.7	398.8	176.4	214.9	302.1
	(206.4)	(252.9)	(290.0)	(187.6)	(232.0)	(267.0)
Mean Square Error – R.D.	217.6	260.7	296.4	132.7	174.8	231.6
	(211.7)	(244.9)	(289.4)	(196.1)	(224.2)	(264.3)
Coverage Error Rate –R.D.	247.2	304.8	404.6	180.1	234.3	341.1
	(240.2)	(278.6)	(317.1)	(222.4)	(254.6)	(286.7)

Note: Dependent variable is the number of observations for each birth month cohort. The top panel reports the coefficient of D when the optimal bandwidth method used is Mean Square Error – Two Sided (two different optimal bandwidth choices above and below the threshold). The bottom panel reports the same coefficient for other optimal bandwidth choices, for robustness check. Robust-heteroskedastic standard errors in parentheses.
 *** p<0.01, ** p<0.05, * p<0.10
† Optimal bandwidth choice method is Mean Square Error – Two Sided (two different optimal bandwidth choices above and below the threshold).

Table A3 – Returns to College Education for Men Using the Discontinuity from the Cultural Revolution as an Instrument

	Rural				Urban			
	OLS	IV			OLS	IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College & above	0.979***	1.134***	1.056***	1.045***	0.675***	1.149***	1.142***	1.139***
	(0.040)	(0.181)	(0.155)	(0.156)	(0.015)	(0.057)	(0.056)	(0.056)
Age	0.079***	0.078***	0.078***	0.078***	0.122***	0.112***	0.112***	0.112***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.006)	(0.006)	(0.006)
$Age^2 \times 10^{-3}$	-0.876***	-0.864***	-0.870***	-0.871***	-1.336***	-1.196***	-1.198***	-1.199***
	(0.047)	(0.050)	(0.048)	(0.048)	(0.075)	(0.064)	(0.064)	(0.064)
Constant	11.877***	11.874***	11.876***	11.876***	11.225***	11.310***	11.309***	11.308***
	(0.092)	(0.091)	(0.091)	(0.090)	(0.138)	(0.129)	(0.129)	(0.128)
Polynomial order of the RD in the 1st Stage		1st	2nd	3rd		1st	2nd	3rd
F-Statistic[†]		57.5	116.6	106.6		285.9	271.4	291.2
Observations	49,615	49,615	49,615	49,615	60,377	60,377	60,377	60,377

Note: Dependent variable is the natural log of average wage for every job category, calculated from the Household Expenditure and Income Survey of 2010. The sample for all these regressions consists of birth cohorts between 1950 and 1979 (i.e. ±15 years around the threshold.)
*** p<0.01, ** p<0.05, * p<0.10
† Kleibergen-Paap rk Wald F statistic for weak instruments.

Appendix B

We need a more careful consideration of data than a visual observation to see if there has been a statistically significant difference in the educational attainment of the affected and partially affected cohorts relative to others. One way to do this is to look at the deviation from the trend for these cohorts. In other words, one may run the following regression:

$$C_{is} = \alpha_s + \beta_1 T1_{is} + \beta_2 T2_{is} + f(B_{is}) + u_{is} \qquad \text{(B1)}$$

in which C_{is} is a dummy equal to one if individual i in region s has partially or fully attended undergraduate or graduate schools. α_s are region fixed effects. There are two regions per province: one representing all the urban areas in a province and another for all the rural areas in that province. Because there were 30 provinces in 2006, the total number of clusters is 60.[a] All results have robust-heteroskedastic standard errors corrected for within region correlations. $T1_{is}$ is a dummy variable equal to 1 if individual i in region s belongs to birth cohorts that were fully affected (i.e. 1962 through 64) and zero otherwise. $T2_{is}$ is a dummy variable equal to 1 if individual i in region s belongs to birth cohorts that were partially affected (i.e. 1959 through 61) and zero otherwise. B_{is} is the year individual i in region s was born (i.e. it is the birth cohort). $f(B_{is})$ is a function that picks up a linear or non-linear trend. Based on the shape of the graphs in Figures 1(a) and 1(b), we use 1) a linear trend (only B_{is}), 2) quadratic trends (B_{is} and B_{is} squared), and 3) differential linear trends before and after the Cultural revolution. β_1 and β_2 show deviations from the trend for cohorts who were fully or partially affected by the Cultural Revolution. Table B1 reports the results.

---

[a] There were 31 provinces in 2011. So, the provinces in 2011 were matched with those in 2006 to create a consistent regional data over time.

Column (1) in Table B1 depicts the regression in Equation (B1) using the sample of males in birth cohorts of 1931 through 1984 and a linear trend. The coefficient of T1 shows that cohorts of men who were fully affected by the Cultural Revolution had about 1.2 percentage points lower rate of college (and above) attainment than the trend. Coefficient of T2 depicts that for those who were partially affected, the attainment rate was the same as the linear trend.

Considering a linear trend for college attainment rate may be strong (see Figures 1(a) and 1(b)). One can consider a quadratic trend. But as Column (2) of Table B1 shows, adding the trend squared, i.e. B_{is}^2, has little impact on the results for the sample of men.

One may argue that after the Cultural Revolution, the transformation of institutions of higher education might have affected the trend altogether. This is easy to control by adding AR_{is} and its interaction with the trend variable, B_{is}, to Equation (B1). AR_{is} is a dummy equal to one if individual i in region s was born after 1961 and zero otherwise. The interaction of this variable and the trend picks up the change in trend after the Cultural Revolution. The results for men are reported in Column (3). Coefficient of T1 shows a -1.5 percentage points deviation from the trend before the Cultural Revolution. But the deviation for the partially affected cohorts remains insignificant and small.

Column (4) represents the same regression as Column (1) but for women. The coefficients of T1 and T2 imply that attainment rates for cohorts fully or partially affected by the Cultural Revolution were 4 percentage points below the trend. But, as shown in Figure 1(b), the trend for women is more like a quadratic one. Hence, squared of birth year should be added to the specification. Column (5) reports the results for such a specification. The deviations found in Column (4) become quite smaller (from about four to less than one percentage points).

Another specification is to control for separate trends before and after the Cultural Revolution. Column (6) depicts the results. The coefficients imply that attainment rate for female cohorts who were fully affected by the Cultural Revolution was 2.2 percentage points *larger* than the trend. This is an odd result and clearly shows that the deviations from the trend for women are sensitive to the specification.

Table B1 – Deviations from the Trend for College Attendance Because of Cultural Revolution

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
T1	-0.012***	-0.013***	-0.015***	-0.043***	-0.009***	0.022***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)
T2	-0.003	-0.005	-0.005	-0.039***	-0.004**	-0.001
	(0.004)	(0.003)	(0.006)	(0.003)	(0.002)	(0.004)
Birth year	0.003***	0.004***	0.003***	0.005***	-0.007***	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Birth year^2 / 1000		-0.006			0.144***	
		(0.007)			(0.008)	
AR			0.019			-0.373***
			(0.025)			(0.023)
Birth year × AR			-0.0002			0.008***
			(0.0002)			(0.000)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	734,183	734,183	734,183	736,074	736,074	736,074
Number of Regions	60	60	60	60	60	60

Note: Dependent variable is a dummy equal to one if the individual partially or fully completed college and above levels, and zero otherwise. Robust-heteroskedastic standard errors corrected for correlation within regions in parentheses. A Region is all urban or rural areas within a province. The sample consists of birth cohorts between 1931 and 1981.
*** $p<0.01$, ** $p<0.05$, * $p<0.10$